DISSERTATION

MULTISCALE CONNECTIONS BETWEEN A GROUNDWATER DEPENDENT ECOSYSTEM AND SOCIO-HYDROLOGY: INSIGHT GAINED FROM NUMERICAL MODELING, GEOSPATIAL INFORMATICS, AND BAYESIAN STATISTICS

Submitted by

Matthew R. Lurtz

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2023

Doctoral Committee:

Advisor: Ryan R. Morrison

Aditi S. Bhaskar
Ryan T. Bailey
Matthew Ross

ABSTRACT

MULTISCALE CONNECTIONS BETWEEN A GROUNDWATER DEPENDENT ECOSYSTEM AND SOCIO-HYDROLOGY: INSIGHT GAINED FROM NUMERICAL MODELING, GEOSPATIAL INFORMATICS, AND BAYESIAN STATISTICS

The connectivity between floodplain practices and groundwater dependent ecosystems (GDE) is undeniable, yet difficult to measure. Quantifying the connection between ecosystems would be ideal for the conjunctive management of groundwater and surface water resources in an irrigated river valley. In the research presented, a variety of methodologies are used to understand the socio-hydrologic connections between a semi-arid GDE and agro-pastoral practices in southeastern Colorado (USA).

I investigated the socio-hydrologic relationships between a GDE and the surrounding floodplain using three approaches. First, I used the output from a calibrated groundwater model and a remote sensing evapotranspiration (ET) algorithm with exploratory statistics. Second, I used remotely sensed vegetation information and socio-hydrologic data in a Bayesian hierarchical time series and spatial statistics models to compliment the first approach by examining new explanatory covariates. Third, a simple regression framework examines the point-scale relationship between groundwater and ET to further dissect results from the first approach at a finer resolution. These three approaches yielded key results.

From the first objective, the dual-model comparison agreed with previous ecological research showing a non-linear relationship between ET and groundwater depth (0-5 m), and a threshold was identified at three meters where the rate between ET and groundwater depth change. The time series and spatial statistics objective helped identify a spatial scale threshold to detect temporal trend, lagged intra-seasonal predictors of vegetation water use, and which floodplain characteristics impact vegetation density. This statistical analysis discovered that temporal trend is not detectable

at spatial scales larger than catchment size ($>$ 10 km). Monthly temperature and lagged monthly values of precipitation and stream gain-loss (i.e., an return flow indicator variable) are all predictive of temporal changes in riparian vegetation density. Based on the floodplain characteristics tested in the spatial statistics approach, perennial tributaries to the Arkansas River increase vegetation density while the conversion of agriculture to fallow land decrease riparian vegetation density. The third objective highlighted that the process between evapotranspiration and groundwater head is non-linear but depends on temporal scale and plant functional group. The results from these approaches is important for GDE preservation in the face of increasing demand on groundwater supply. The process between groundwater and ET is of particular importance in large scale water balance studies that include a groundwater and surface water interface with need to model the groundwater-ET relationship in natural and agricultural ecosystems simultaneously.

ACKNOWLEDGEMENTS

I would like to thank my mother (Cheryll Lurtz), my advisor (Ryan Morrison), and my committee (Aditi Bhaskar, Ryan Bailey, and Matthew Ross). Also, a special thanks to Gabriel Senay, Timothy Gates, David Ketchum, Pamela Nagler, Jeff Penn, Mevin Hooten, Thompson Hobbs, Tom Siller for providing helpful comments, technical assistance and constructive advice.

Maktub!

DEDICATION

I would like to dedicate this thesis to the late Dr. Jorge A. Ramirez because without his belief in me, the next chapter of my life as a National Science Foundation Postdoctoral Research Fellow would not be as fulfilling.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

General Introduction

1.1 Motivation

In the Earth's hydrosphere, competition for beneficial uses (e.g., agriculture, energy, and ecosystem services) is recognized as probably the largest issue facing humanity in a changing climate [1]. There are numerous rivers in the world categorized as Endangered Rivers that supply water to millions of people [2]. These river-aquifer-ecologic systems are rigidly connected and are under serious threat from groundwater extraction, pollution and climate change [3]. Recently, evidence of spatial patterns of stream flow losses have been observed in the contiguous United States [4]. The disappearance of surface water supplies will accelerate the need for groundwater extraction to meet future demands which will lead to dysfunction in natural ecosystem services. Groundwater resources are therefore under threat of development, and there is an urgent need to understand how groundwater supply is linked to groundwater dependent terrestrial ecosystems, and how the health of these ecosystems are controlled by the resources of Endangered Rivers.

1.2 Background

A background understanding is presented herein to highlight key concepts of this research. A commonality between Chapters 2, 3, and 4 is quantitatively describing the connection between a groundwater dependent ecosystem (GDE) and the surrounding floodplain in southeastern Colorado. Therefore, background on GDEs in arid climates and the socio-hydrologic connections between floodplains and GDEs is described.

Groundwater dependent ecosystems are ecosystems that depend on groundwater to satisfy a percentage of their water use requirements [5,6]. Groundwater dependent ecosystems span a range of types such as springs, karst systems, rivers and lakes and can be found in coastal and inland environments. In this research, I focus on low gradient inland GDEs that are also described as

riparian ecosystems. GDEs are well known and have been studied in China [7], Australia [3, 8, 9], Europe [10, 11], and across the western United States [5, 12–14]. In the arid western part of the United States, GDEs can play a key role in sustainably managing interconnected groundwater and surface water systems in irrigated agricultural regions. Vegetation water use, derived from groundwater or another source, can be quantified using field data, numerical modeling, and satellite imagery [15–18]. More specifically, GDE evapotranspiration, an indicator of ecosystem function, is commonly measured using satellite-derived reflectances in unison with gridded and point based micrometeorological data [19–21]. Evapotranspiration from a GDE that is positioned between a river and an agroecosystem has inherent socio-hydrologic connections.

The main pathway by which there are socio-hydrologic connections to a groundwater dependent ecosystem is described herein. Historically, alluvial river valleys would irrigate crops using flood and furrow techniques in which excess water, not consumed by the crop or evaporated, would infiltrate the soil and recharge groundwater [22]. This irrigation-induced recharge will move through the aquifer and eventually feed a surface-water body [23]. To reduce the amount of water applied for the same crop yield, advanced irrigation technologies have been adopted (flood-furrow to center-pivot) which can reduce groundwater recharge and alter groundwater availability for riparian vegetation [24]. In many Colorado watersheds, especially within the Lower Arkansas River watershed in southeastern Colorado, irrigated agriculture is the most significant water user [25]. The connection between floodplain activity and riparian flora through return flow is an area of hydrologic science that deserves increased attention. Because GDEs have complex spatio-temporal dependence structures when compared to agroecosystems, these ecosystems are often neglected in large-scale water-balance studies [26]. There is need to further understand the relationship between floodplain activity and GDEs because they are connected through hydrologic exchange flows which are of primary concern in water conservation strategies in irrigated alluvial valleys [25, 27].

This research narrows the knowledge gap found in GDE and groundwater-agriculture research. Comprehensive GDE literature and key research on irrigated alluvial valleys in Colorado and California present gaps that this work aims to advance. Specifically, there are two key reviews that

summarize new research directions for GDEs. A review into GDE water-vegetation interactions concluded that among several topics to advance GDE research, knowledge gaps include how water is used by different vegetation types and the role of GDEs in floodplain management [19]. Even further, understanding how hydrologic exchange flows impact different vegetation communities is an understudied area. Another review indicated that examining vegetation adaptation to differences in water supply are worthy directions for GDE research [28]. Regarding irrigated alluvial valleys, Foglia et al. (2013) highlights the importance of field monitoring efforts to better understand site-specific groundwater-stream interactions [29]. Because of these gaps, the research I performed focuses on vegetation water use in GDEs and how the patterns of water use are connected to the surrounding floodplain. The research performed in Chapters 2, 3, and 4 aim to fill these knowledge gaps using the Arkansas River in southeastern Colorado as a primary study area.

1.3 Overview of Chapters

The research presented herein focuses on groundwater dependent ecosystems through two major lenses: 1) the relationship between evapotranspiration and groundwater depth, and 2) the socio-hydrologic connections between an agroecosystem and a riparian ecosystem. This research is divided into three approaches that differ upon scale. Common hydrologic definitions for scale as proposed by Blöschl and Sivapalan (1995) are adopted [30]. The three approaches are summarized in the following bulleted list:

- Catchment-scale (10 km) relationships between riparian evapotranspiration and groundwater are analyzed in time and space using output from a finite difference groundwater model and a satellite based evapotranspiration algorithm.

- Reach-scale (100 m) relationships between riparian vegetation and socio-hydrologic characteristics are quantified using hydrologic data and satellite imagery and are examined using independent, autoregressive space and time Bayesian linear models.

- Point-scale (1 m) function of best fit is derived between normalized evapotranspiration and groundwater using a Bayesian regression framework.

The drivers of GDE water use should be examined further. Specifically, the three scientific approaches used in this research will help water and land managers see the influence that agroecosystems have on GDEs. This type of research can improve management implementation of conjunctive use schemes that include the interface between groundwater and surface water [29]. This research provides ecosystem-variable thresholds important for ecosystem conservation, temporal and spatial statistics that describe the groundwater-agriculture nexus, and how the mechanistic process between evapotranspiration and groundwater varies upon vegetation type.

Chapter 2

Catchment-Scale Relationship Between Riparian Evapotranspiration and Groundwater

2.1 Introduction

The revolving relationship between anthropogenically-influenced groundwater depth and riparian evapotranspiration (ET), especially in arid climates, is a unique aspect of ecohydrology that entangles social dimensions and natural-ecosystem research [31]. Riparian flora supports stream and sediment stabilization [32], ecosystem function [33], and removal of agricultural contaminants [34, 35]. Quantifiable relationships between vegetation and water table depth, D_{wt}, can inform current and future numerical models [36] that include riparian-ecosystem complexity. For example, the U.S. Geological Survey's (USGS) modular hydrologic model MODFLOW employs a traditional linear response of ET to groundwater head while the Evapotranspiration Segments Package and the Riparian Evapotranspiration Package provide a more complex way of simulating ET in MODFLOW as a function of groundwater supply [37]. Riparian ET remains an elusive component of water-balance studies due to its complex connection with groundwater at various time scales ranging from daily hyporheic-zone fluctuations to annually-consistent regional flow [19], which allow the vegetation to optimize its potential for success by adapting to groundwater levels [28]. With the help of independent data sets from a calibrated finite-difference model simulation and validated remote-sensing methods, additional insight into modeling riparian ET at varying spatio-temporal scales can provide information to improve the body of knowledge surrounding groundwater-dependent ecosystems [28].

There are numerous models for efficient and reliable estimation of ET using remotely-sensed data, such as the Surface Energy Balance for Land (SEBAL) [38], Remote Sensing of ET (ReSET) [39], and Mapping ET at high Resolution with Internalized Calibration (METRIC) [16],

which have focused on cropped landscapes. Natural ecosystem ET has also been investigated using several techniques [40–43] that vary in input data, spatial resolution, and recommended applicability. Although all of the latest remotely-sensed ET models are unique, they each are inherently more or less a self-operating scatterplot method. Included in the list of validated models applied to agriculture is the Operational Simplified Surface Energy Balance model (SSEBop) [44] which is an automated method to produce ET on the day of satellite overpass. The SSEBop model approaches the surface energy balance in a simplistic way relative to SEBAL and METRIC's algorithms, yet it is appropriately generalized to apply in natural ecosystems adjacent to agricultural areas. The SSEBop algorithm uses changes in land surface temperature and reference ET as the model forcings that support changes in ET. The validated model operates under the premise that colder land surface temperatures represent areas of increased ET relative to warmer surface temperatures, much like an automated-intelligent scatterplot method [45]. Presently, the SSEBop approach has been applied using varying temporal scales with different Landsat sensors, which have the spatio-temporal resolution required to capture riparian lifecycles [46, 47].

Previous studies have estimated remotely-sensed riparian water use along the Colorado River [48, 49], the Rio Grande River [42], the Verde River [50], and the San Pedro River [51], all of which supply water to irrigation districts with economic significance. The Arkansas River is an equally important water source that suffers from a similar non-native species invasion of riparian ground but lacks a competing amount of ET estimates [26].

Similar to many irrigated basins, groundwater levels in the Arkansas River valley, including within the riparian ecosystem along the river, are influenced by agricultural irrigation. The size and timing of return flows can be modified both by varying hydrologic conditions and by watershed managers in areas with excess or inefficient irrigation practices [52], and changes in return flows could subsequently impact riparian water use. Irrigation-influenced return flows can provide a considerable amount of water to surface water bodies, and studies have been performed to quantify return flows [22,53,54] and patterns of return flows using data-supported computational modelling [52, 55]. A groundwater model, which utilizes MODFLOW, developed by Morway et al. (2013)

and refined by Shultz et al. (2018a) is employed in this study. The MODFLOW model was developed to investigate the dependencies of regional water table elevations and groundwater flow to management decisions aimed at water conservation, agricultural production, and maintaining aquatic ecosystems. Riparian ET, root-extinction depth, and agricultural return flows are highly researched topics, and connecting them using a calibrated groundwater model and remotely-sensed ET can inform how floodplain activity impacts the adjacent ecosystem.

For the research presented, I have seized a unique opportunity to investigate the temporal relationships between datasets of groundwater levels from a finite-difference groundwater model and ET calculations from a validated remote-sensing ET algorithm. The main objective of this study is to statistically compare D_{wt} from a highly-calibrated regional groundwater model, which spans over a decade of hydrologically variable years, to the magnitude of SSEBop-calculated ET in areas with different groundwater-fluctuating patterns. Investigations of this scale will further the understanding of complex interactions between riparian water use and groundwater as influenced by both natural hydrology and water management.

2.2 Methods

2.2.1 Study Site

The riparian corridor targeted in this study lies within Colorado's Lower Arkansas River Valley (LARV), extending eastward of Manzanola, Colorado (CO) (Figure 2.1), following the Arkansas River to the Adobe Creek tributary near Las Animas, CO. The LARV is semi-arid where evaporation exceeds precipitation as defined by the Ecoregions of North America map [56]. The areal coverage of the riparian corridor within the study area, including the surface area of the Arkansas River, is 38.5 km^2, with good hydraulic connection between the alluvium and the river [52]. The Colorado Climate Center manages micrometeorological data collected at Colorado Agricultural Meteorological (CoAgMet) (CoAgMet) stations throughout the state. The nearest CoAgMet station (RFD01) is located 4.0 km southeast of Rocky Ford at the Colorado State University (CSU) Arkansas Valley Research Center (AVRC), and began collecting data in 1992. The RFD01 sta-

tion data are utilized for the analysis given the proximity to the riparian corridor and the longevity in data collection. As recorded by the RFD01 station using a Campbell Scientific tipping-bucket data logger, the 11-year average annual rainfall in nearby Rocky Ford, CO, from 1999 to 2009 is 249 mm. The lowest annual rainfall was recorded in 2002 at 72 mm and the highest was in 2004 at 352 mm. During the period of analysis, April 1999 – October 2009, the highest one-month rainfall record was recorded in August 2008 at 110 mm and in April 1999 at 103 mm. Parameter-elevation Regressions on Independent Slopes Model (PRISM) data, along with data from the National Weather Service Cooperative Observer Network (COOP), are compiled into an interactive online database, Westwide Drought Tracker, which indicates that the Upper Arkansas Hydrologic Unit (HUC: 11020002), where the study area resides, experienced above normal annual rainfall totals in 1999, 2000, 2005, and 2007 (https://wrcc.dri.edu/wwdt/). Westwide Drought Tracker indicated below average conditions for the other seven years of the study period, including the years 2002 and 2003 which were two of the driest years on record for the southeast region of Colorado and much of the state. Normal conditions are based on monthly averages estimated for the HUC since 1895 to present. Because rainfall is minimal, surface water diversions and pumped groundwater are used to support agricultural production. Irrigation in the LARV is practiced mostly by flood and furrow, with a smaller percentage of fields using sprinklers and drip lines [57].

The riparian corridor, within the study region, extends approximately 70-km along the Arkansas River, as shown in Figure 2.1. In this semi-arid region of southeastern Colorado, extents of the riparian corridor are determined with visual inspection of satellite imagery for vegetation and changes in surficial geology. The river's edge is delineated by visual inspection of the green-line width [58]. Historical cropping patterns in the LARV were delineated from naturally-vegetated areas using the National Land Cover Database 2004 dataset, downloaded from the Multi-Resolution Land Characteristics Consortium (https://www.mrlc.gov/), with knowledge of past irrigation-district extent. Per the NLCD-2004 data set, the riparian corridor is comprised of about 47 percent emergent herbaceous wetlands, 17 percent woody wetlands, 14 percent grassland/herbaceous, 11 percent open water, 6 percent shrub/scrub, with the remaining 5 percent disbursed among land classes

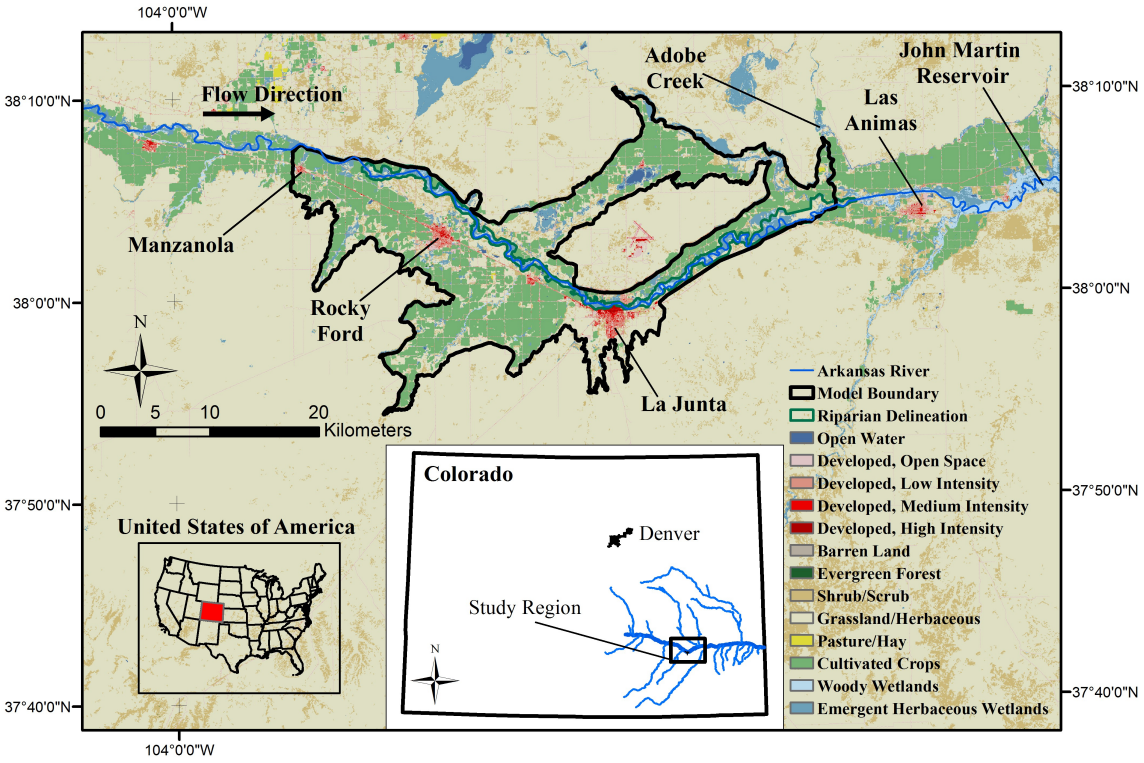

Figure 2.1: Location map of the study area including groundwater model boundary, riparian delineation, and an outline of the Arkansas River.

overlaying canals and livestock practices. Historically, the riparian flora along the Arkansas River were comprised of various types of river birch, shrublands and cottonwoods [59]. More recently, Colorado's Lower Arkansas River basin has been impacted by the spread of salt cedar and Russian olive [59].

2.2.2 Groundwater Data and Analysis

The MODFLOW model employed in this study overlaps 506 square kilometers of land area and employs 15,600 active computational grid cells with areal dimensions 250 m x 250 m and two vertical layers. The model simulates 574 weeks of unconfined groundwater level and flow in three dimensions using a finite-difference approximation of the governing equations of unconfined aquifer flow. The model incorporates the UZF module for computing unsaturated-zone water content and vertical flow using a simplification of the Richard's equation and an estimation of upflux from shallow groundwater as a function of D_{wt}, for the time frame April 1999 through October

2009. Observations from approximately 100 monitoring wells and estimates of return flows from a river water balance were used, along with other data sets, to calibrate and test the model. The simulated groundwater data used in this study consisted of weekly hydraulic head values from 1999-2009. The weekly output of hydraulic head was aggregated to a weighted-average monthly value per model cell. Using the number of weeks of output that occur within a month, typically five or six, each weekly matrix of hydraulic head was multiplied by the number of days it represented in a particular month (1-7 days) and the sum of the weekly scaled values were divided by the total number of days within a given month. Monthly values of D_{wt} were then computed as the weighted-average hydraulic head subtracted from the land surface elevation used in the groundwater model. The groundwater model data used in this study to generate average monthly D_{wt} values are presented in Morway et al. (2013), Gates et al. (2016), and Shultz et al. (2018a), which describe in detail the groundwater model built for the LARV and the supporting field data, including crop and irrigation data. Readers can refer to Morway et al. (2013) and Shultz et al. (2018a, 2018b) for specific information on the model's development, testing, and application.

2.2.3 Other Datasets

The application of SSEBop in this study required numerous geo-processing steps and subsequent testing with independent data sets. The annual numbers of satellite images used to calculate SSEBop ET are shown in Table 2.1. Before using the satellite images in the SSEBop algorithm, each image was inspected for cloud contamination using the USGS Earth Explorer website (http://earthexplorer.usgs.gov/). Satellite images used in the analysis had cloud cover less than 10 percent and went through a secondary filter to eliminate those not fit for the ET algorithm due to either cloud or ice coverage, as described in detail in Senay et al. (2017). The Landsat 7 scan line corrector failed in 2003; hence, the corresponding missing values are left as not available (NA) on the day of the satellite pass and removed from any further analysis. The cell location found in error was not used to calculate monthly sums of SSEBop ET or for any subsequent analysis after. Land surface temperature (Ts) was computed using band 6 from Landsat

5 and 7 data following the same methodology outlined in Senay et al. (2016). Two gridded inputs to the SSEBop algorithm are the daily alfalfa reference ET (ETr), available from gridMET (http://www.climatologylab.org/gridmet.html), and daily maximum and minimum air temperature (Ta), used directly in the computation of net radiation (Rn) and air density (ρ_a), available from TopoWx (https://www.ntsg.umt.edu/project/topowx.php). Alfalfa reference ET from gridMET (gridMET ETr), is calculated using the ASCE Penman-Montieth method (ASCE-PM) [60]. TopoWx provides gridded temperature with high spatial resolution across the conterminous United States based on numerous validated data sets from station observations and satellite data.

Table 2.1: Annual number of Landsat images used to calculate SSEBop ET.

Year	Landsat 5	Landsat 7	Total
1999	14	7	21
2000	7	10	17
2001	11	7	18
2002	8	9	17
2003	15	9	24
2004	9	9	18
2005	11	10	21
2006	13	9	22
2007	9	11	20
2008	9	15	24
2009	13	9	22
Total	119	105	224

Estimates of ET calculated using micrometeorological station data, along with local measurements of ET using a nearby weighing lysimeter, were compared to SSEBop ET estimates. The COAgMet RFD01 station data are utilized for the analysis. Data downloaded from RFD01 consisted of daily totals and averages of ETr, rainfall, solar radiation and mean air temperature. The standardized ET equation used by CoAgMet for estimating a ETr is also the ASCE-PM equation. Al Wahaibi (2011) tested the ASCE-PM ETr equation against precise measurement of the ET of

alfalfa in a weighing lysimeter (3 m x 3 m x 2.4 m) at the CSU AVRC. The weighing lysimeter began operating in 2007 and alfalfa was harvested four times in the 2008 and 2009 growing seasons. As reported in Al Wahaibi (2011), deficit soil moisture conditions were experienced for the entire 2008 growing season and until June 8, 2009 (DOY 159) with average daily lysimeter ETr for 2009 at 6.0 mm/day. Lysimeter estimated ET is referred to as Lysimeter ETr in this analysis because it represents fully-irrigated alfalfa water use that was measured on an hourly basis by precisely monitoring the change in mass, with daily values calculated from a water balance accounting of precipitation, irrigation, and change in storage. Drainage is also a component of the water balance but is negligible during the period of study [61]. Using the aforementioned lysimeter, the lysimeter construction and previous experimentation has been described in literature [62, 63].

2.2.4 SSEBop ET Calculations

The ET model used in this study is an operationalized version of the original SSEB model [64] that circumvents the issue of requiring the model user to select "hot" and "cold" pixels. This enhancement of the SSEBop algorithm is synonymous with removing manual selection of physical boundaries represented by scatterplot methods, in which ranges of a vegetation index and corresponding land-surface temperature, per pixel, are plotted in an x-y scatterplot fashion to display the biophysical gradient (e.g. "hot" and "cold" cells) across a given satellite scene. Scatterplot methods are explained in detail by Petropoulos et al. (2009). Previous literature has justified the use of "hot" and "cold" cell remote-sensing applications in semi-arid regions, allowing for the implementation of SSEBop in southeastern Colorado to be an appropriate experiment to build knowledge on natural ecosystem water use [65]. The application of SSEBop in this research closely follows that of previous efforts by Senay et al. (2013, 2016, 2017). Although an overview of the SSEBop model is provided here, readers are referred to the above references for further explanation. The SSEBop model is driven by available net radiation and advective forces that determine ET_r, and the differences in land surface temperature are intended to capture the range of ET occurring over a land surface. Unlike most applications of SSEBop, alfalfa ET_r is used here to better represent the

rougher vegetation found in riparian corridors, thus precluding the need for a scaling k factor:

$$\text{SSEBop ET} = \text{ET}_f \times \text{ET}_r, \tag{2.1}$$

ET_f in Eq. (2.1) is the ET fraction calculated as (2.1)

$$\text{ET}_f = \frac{\text{T}_h - \text{T}_s}{\Delta\text{T}} = \frac{\text{T}_c + \Delta\text{T} - \text{T}_s}{\Delta\text{T}} = 1 - \frac{\text{T}_s - \text{T}_c}{\Delta\text{T}} \tag{2.2}$$

where T_s (^{o}K) is the satellite-observed land surface temperature of a pixel; T_h (^{o}K) is the estimated T_s at the idealized "hot" pixel boundary; and T_c (^{o}K) is the estimated T_s at the idealized "cold" boundary. The temperature difference between hot and cold pixels, ΔT (^{o}K), is defined as [44]

$$\Delta\text{T} = \frac{R_n r_{ah}}{\rho_a C_p} \tag{2.3}$$

where R_n is daily average clear-sky radiation (J m^{-2} s^{-1}); r_{ah} is the aerodynamic resistance to heat transfer, chosen as 110 s m^{-1} to remain consistent with past SSEBop implementations that track relative changes in ET [44]; ρ_a is the density of air (kg m^{-3}); and Cp is the specific heat capacity of air at constant pressure (1.013 kJ kg^{-1} K^{-1}). Consistent with Senay et al. (2016), this study uses a value of 1.05 as the maximum for ETf, negative ETf values are set to zero, and cloud-contaminated pixels are set to "no-data".

The value of T_c at cold boundary pixels is calculated using the methodology found in Senay et al. (2017) and are derived from T_a. The model assumes that for well-watered vegetation, with ET occurring at a rate no less than allowed by micrometeorology, T_c will equal T_a. The factor c = T_s/T_a is uniquely defined for each Landsat image, following an image-filtering process shown in Senay et al. (2017), to highlight the wet/cold cells in the satellite image. In summary, the image-filtering process chooses pixels with increased potential to evapotranspire, relative to other pixels within a Landsat scene, by using the normalized difference vegetation index (NDVI) as a "greenness" indicator and by removing cells with unrealistic temperature differences, usually

caused by clouds or snow. When an image fails to meet the conditions required to develop a c factor correction, the satellite scene is not used in the analysis. In order to screen each Landsat image for the presence of clouds and cloud shadows, the Function of Mask algorithm is applied in a programmed manner [66]. Datasets integral to the SSEBop algorithm and post-processing analysis are summarized in Table 2.2.

Table 2.2: Summary description of data sets used in this study.

No.	Source	Variables	Frequency	Resolution
1	Landsat 5	Surface Temp. and NDVI	16 days	30 m
2	Landsat 7	Surface Temp. and NDVI	16 days	30 m
3	TopoWx	Air Temp. and Net Radiation	Daily	800 m
4	Mapzen	Elevation	–	30 m
5	gridMET	Ref. ET	Daily	4 km
6	CoAgMet	Precip., Alfalfa ETr, Air Temp., Net Rad.	Daily	Point
7	Al Wahaibi (2011)	Lysimeter ET	Daily	3 m $\times$ 3 m

Our area of study lies completely within a single Landsat image (path 32, row 34); therefore, SSEBop processing did not require creating mosaics from multiple scenes. Computed SSEBop ET pixels overlaying the Arkansas River were removed from the analysis portion of the study as this analysis is not concerned with changes in surface water evaporation. Landsat images have a frequency of 16 days between satellite overpass. Including two satellite sensors for the SSEBop application allows the satellite overpass frequency to increase to every eight days, which usually takes place between May and September. The SSEBop model provides daily ET on the days of satellite overpass with a spatial resolution of 30 m x 30 m. Daily SSEBop ET can be summed to any cumulative value by linearly interpolating the ETf value between two satellite image dates. Linear interpolation yields a specific ETf value, which can be multiplied by the daily ETr to produce SSEBop ET for any day of the year. This procedure follows the methodology in Senay et al. (2017).

2.2.5 Analysis of ET-groundwater depth relationships

For this investigation, daily values of SSEBop ET are summed to monthly totals and compared to average-monthly computed D_{wt}, unless explicitly noted. Monthly sums of SSEBop ET were chosen as the best temporal discretization to relate to the groundwater-model output for two reasons: 1) the finest temporal discretization from the MODFLOW model is weekly and 2) estimations of SSEBop ET are typically reported as summations beyond daily time scales to decrease systematic errors. Spatially, there are more simulated estimates of SSEBop ET when compared to the corresponding values of D_{wt}. A 1:1 relationship between pixel-based SSEBop ET and pixel-based D_{wt} is achieved by taking the arithmetic mean of a group of ET-pixel centroids that lie within the spatial extents of a MODFLOW grid cell, all within the areal extent of the riparian corridor with heterogeneous vegetation types. Spatial-averaging is conducted to reduce noise in the ET data set when there is more than one land cover type per 250 m by 250 m MODFLOW grid cell. For the temporal analysis, a spatial average of all values of SSEBop ET is computed to generate one value of riparian ET per month in a given year.

Spearman rank correlation, ρ, was used to investigate the relationship between the groundwater and ET datasets. Spearman ρ was calculated for pairs of SSEBop ET and D_{wt} across 757 grid cells and between the months March–October of the crop growing season in the wet year 1999 and for pairs across 600 cells and over the growing season in the dry year 2003. The difference in pixel counts between years is likely due to the introduction of Landsat 7 into the 2003 analysis when it began having missing cell values. A few values of D_{wt} that were above the land surface elevation (i.e. negative) were assumed to be model computational errors in an attempt to simulate saturated conditions. These values were set to zero before any statistics were calculated. The mid-rank of the sample set was applied to tied ranks [67]. Values of ρ range between -1, which indicates a perfect inverse correlation (e.g. increasing ET with decreasing D_{wt}), to 1, which indicates a perfect direct correlation (e.g. increasing ET with increasing D_{wt}).

Values of ρ also were calculated for each individual MODFLOW grid cell using eight pairs of ET and D_{wt} available for that cell, corresponding to the eight months March–October of the crop

growing season for each of the years 1999 (718 cells) and 2003 (590 cells). The cell-based ρ values lack a large sample size but they inherently contain spatial information about the D_{wt} - SSEBop ET relationship that the monthly estimates do not show. The spatial frequency distribution of these cell-wise ρ values, computed for all of the grid cells that contain a data pair for eight consecutive months, were estimated for the growing seasons of 1999 and 2003. Though statistical significance (α = 0.05) could not be established for the cell-wise ρ values due to the use of only eight monthly-data pairs for each grid cell, they allow for an initial exploration into the degree of spatial variability in the relationship between SSEBop ET and D_{wt}.

2.3 Results

2.3.1 Riparian ET and local micrometeorological conditions

Calculated values of monthly riparian SSEBop ET are found smaller in years with more rainfall, and SSEBop ET values are similar to local micrometeorological station ET data (Figure 2.2). As shown in Figure 2.2, the highest SSEBop ET monthly total of 297 mm occurred in July 2003 with a measured CoAgMET ETr of 289 mm and 12 mm of rainfall. Maximum ETr was observed in June 2002 at 356 mm with a corresponding SSEBop ET of 286 mm and 14 mm of rainfall. Minimum monthly SSEBop ET was computed for December 1999 at 49.7 mm, although minimum CoAgMET ETr occurred in January 2007 at 33.1 mm; but these results were not typical of all years. Wintertime estimates of riparian SSEBop ET are less reliable due to decreased imagery availability and can be summarized as primarily an evaporation estimate given that the riparian vegetation has an annual lifecycle. The largest values of riparian SSEBop ET occurred in June for approximately half of the growing seasons included in the analysis. Values in June 1999 (263 mm), May 2004 (258 mm), July 2007 (256 mm), and July 2009 (249 mm) rank the lowest among peak monthly rates. Average monthly temperature recorded by CoAgMET's RFD01 station shows maximum average monthly Ta (26.8^{o}C) occurred in July 2003, which corresponds to the maximum monthly SSEBop ET computed during the 11-year analysis. Table 2.3 provides a summary of the values for relevant datasets over the entire period of analysis.

Table 2.3: Summary of maximum CoAgMET ET_r with selected climatic variables and months of occurrence; minimum D_{wt} computed by MODFLOW; maximum gridMET ET_r; and maximum spatially averaged riparian SSEBop ET and month of occurrence.

	CoAgMET			MODFLOW	gridMET	SSEBop
Year	ET_r (mm)	Rainfall (mm)	Air Temp. (°C)	Min. D_{wt} (m)	ET_r (mm)	ET (mm)
1999	312.8, Jun.	102.9, Apr.	24.8, Jul.	1.84, May	259, Aug.	263, Jun.
2000	279.3, Jun.	54.9, Mar.	25.1, Jul.	2.13, May	278, Jun.	283, Jun.
2001	296.5, Jun.	92.5, May	25.8, Jul.	3.03, Jan.	285, Jul.	278, Jun.
2002	356.1, Jun.	14.5, Jun.	25.9, Jul.	3.20, Jan.	290, Jun.	286, Jun.
2003	288.9, Jul.	52.6, Apr.	26.8, Jul.	3.20, Jun.	295, Aug.	297, Jul.
2004	267.9, May	91.4, Apr.	22.6, Jul.	3.32, Jun.	275, May	258, May
2005	293.7, Jun.	51.3, Aug.	25.4, Jul.	3.25, Jul	287, Aug.	282, Jul.
2006	301.4, Jun.	92.2, Aug.	25.0, Jul.	3.10, Dec.	278, Jul.	293, Jun.
2007	240.2, Jun.	75.9, Jun.	24.7, Aug.	2.80, Jul.	260, Jul.	257, Jul.
2008	276.4, May	110.5, Jun.	24.7, Jul.	2.48, Nov.	295, Aug.	277, Jun.
2009	237.5, May	67.1, Oct.	23.1, Jul.	1.75, Feb.	275, Jul.	249, Jul.

The weighing lysimeter at the CSU AVRC provides a unique opportunity to evaluate simulated versus measured ET. Daily SSEBop ET estimates, calculated on days of satellite overpass in 2009 for the 100 m x 100 m pixel containing the weighing lysimeter, are compared to daily measurements of alfalfa ETr in the lysimeter, as shown in Figure 2.3. There is a statistically significant (α = 0.05) moderate linear relationship between SSEBop ET and lysimeter ETr (R^2=0.59). Daily SSEBop ET compared to lysimeter ETr in 2009 has a mean absolute error (MAE) of 1.20 mm/d and a root-mean-square error (RMSE) of 1.74 mm/d.

As described earlier, the SSEBop algorithm employs gridMET ETr, estimated using surface meteorological data and interpolated to the nearest 4-km grid, in computing ET using Eq. (2.1) – (2.3). Climatic data from the CoAgMET RFD01 meteorological station are not used as input in the calculation of gridMET ETr. Figure 2.4A shows that spatially-averaged, monthly sums of riparian SSEBop ET have a statistically significant (α = 0.05) strong linear correspondence with CoAgMET ETr (R^2=0.87). Figure 2.4B shows gridMET ETr, compared to CoAgMET ETr, also indicating a statistically significant (α = 0.05) moderate linear correspondence (R^2=0.56). There is

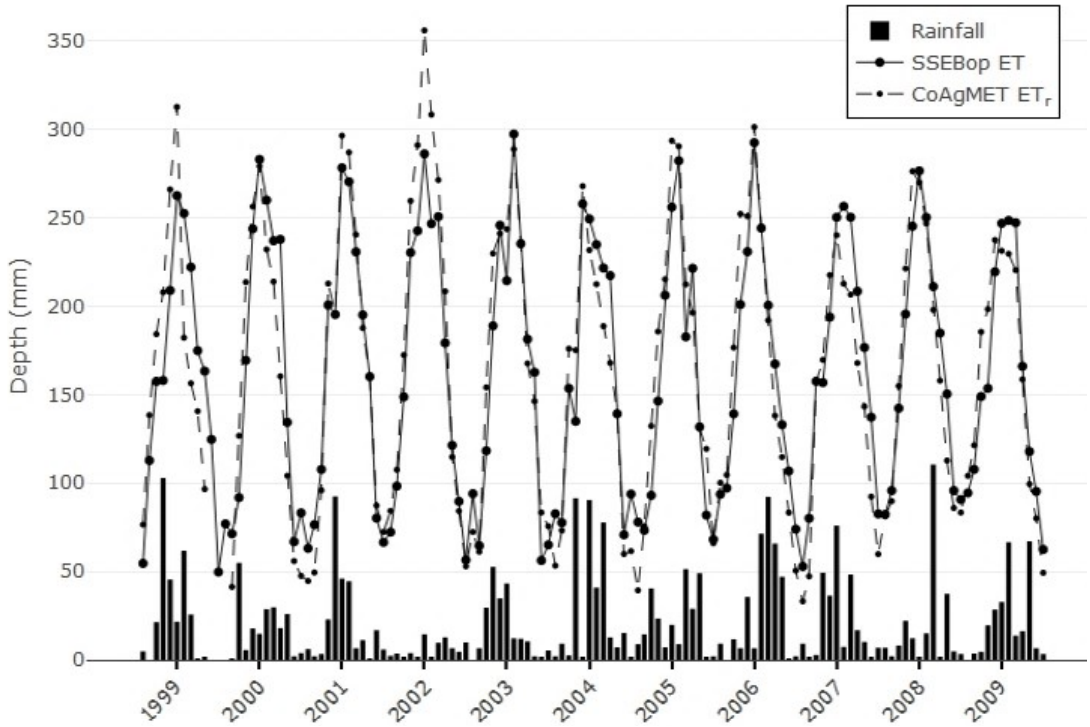

Figure 2.2: Time series representing monthly SSEBop ET, CoAgMET ETr, and rainfall from 1999 to 2009. The riparian SSEBop ET values are spatially averaged, monthly sums. Locally measured rainfall and CoAgMET ETr are monthly sums of daily observations from the CoAgMet RFD01 station. Missing data in November 1999, December 1999, and January 2000 have been excluded from this plot.

statistically significant (α = 0.05) strong linear correspondence (not shown) between riparian SSEBop ET and average monthly atmospheric temperature, Ta, (R^2=0.89) and average monthly solar radiation, Rn, (R^2=0.87). Also, there is statistically significant (α = 0.05) weak linear correspondence with monthly rainfall totals (R^2=0.11); all measured by the RFD01 micrometeorological station. Riparian ET is increased in drier conditions owing to increases in energy demand with heightened temperatures (Table 2.3).

2.3.2 Variations in riparian vegetation, ET, and water table depth

Figure 2.5 illustrates the spatial distribution of average NDVI, peak SSEBop ET, and minimal Dwt during the month of June for wet (year 1999) and dry (year 2003) climatic conditions. Average NDVI values in June 1999 ranged from 0.12 to 0.84 (Figure 2.5A), and in June 2003 ranged from 0.16 to 0.71 (Figure 2.5D). Peak SSEBop ET values ranged from 229 mm to 273 mm in June

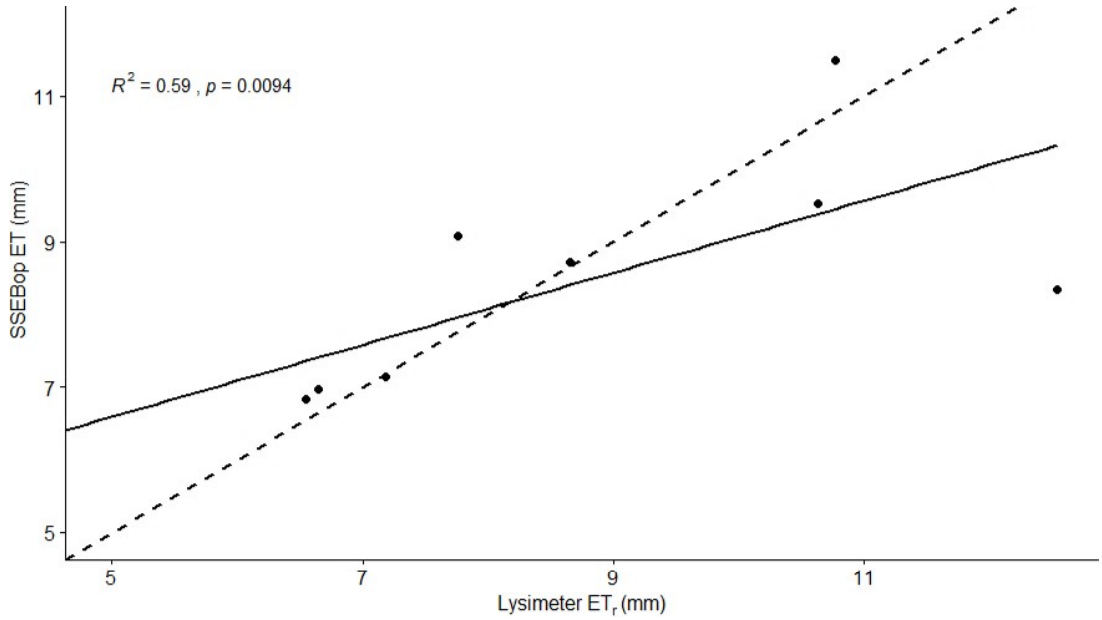

Figure 2.3: A comparison of daily SSEBop ET derived from satellite data to alfalfa ETr measured in the weighing lysimeter at the CSU AVRC in 2009. The least squares linear regression function (solid) and the 1:1 line (dashed) are shown on the graph. Days of satellite overpass prior to June 8 (DOY 159), when the alfalfa crop was minimally irrigated during the first growth cycle, have been removed (n = 10).

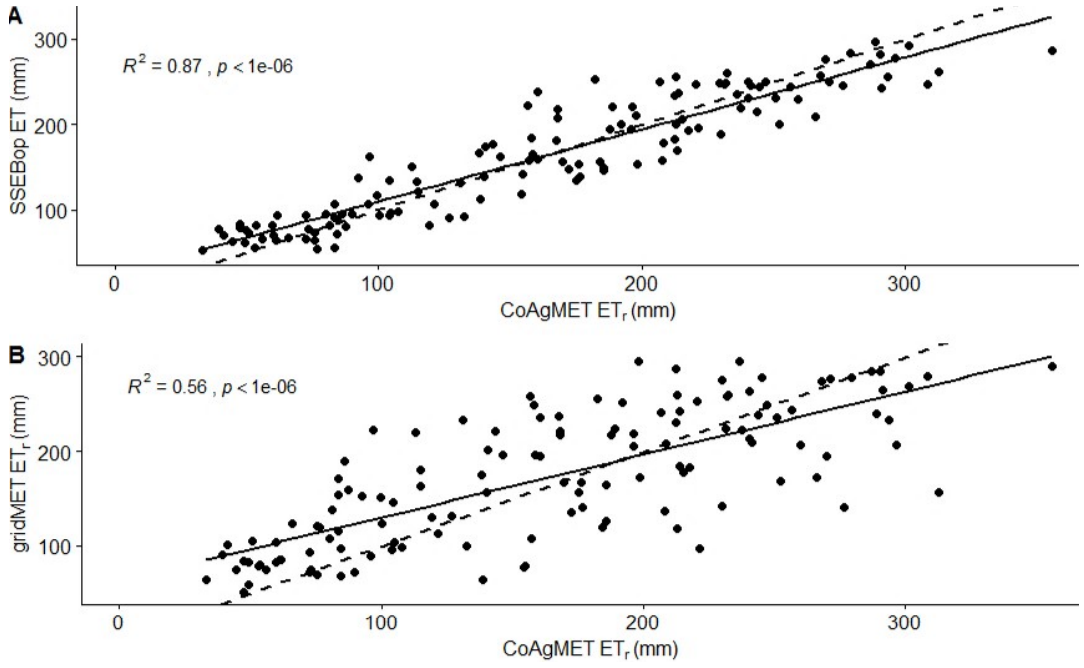

Figure 2.4: Comparison for monthly totals of SSEBop ET (n = 129 data points) against simulated ETr from gridMET and computed ETr from local micrometeorological values collected from CoAgMET's RFD01 station. Data for November 1999, December 1999, and January 2000 are missing. A fitted linear regression (solid) and 1:1 line (dashed) are shown on each graph.

1999 (Figure 2.5B), but were noticeably larger in June 2003 with a range from 228 to 315 mm (Figure 2.5E). Peak SSEBop ET show smaller values in the eastern portion of the modeled area, just upstream of John Martin Reservoir, independent of climatic condition (Figure 2.5B and 2.5E). Larger areas of shallow groundwater depths were present in June 1999 (Figure 2.5C) compared to June 2003 (Figure 2.5F), consistent with the wetter climatic conditions with corresponding higher streamflow and irrigation diversions during 1999. Grid cells with missing data were excluded from Figure 2.5 and were likely due to cloud interference in Landsat images.

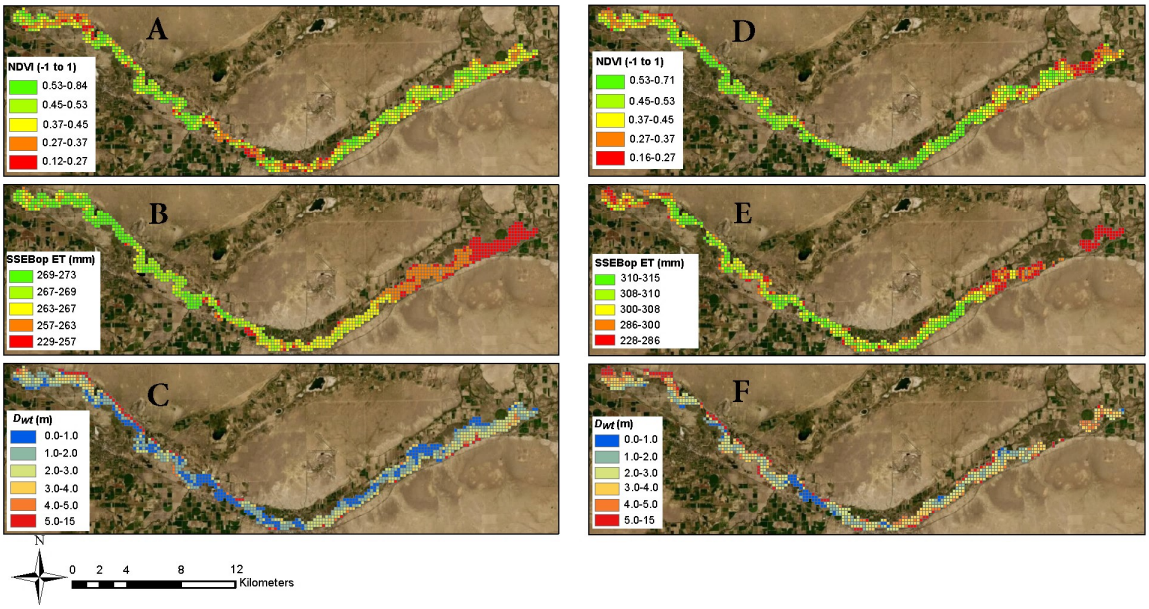

Figure 2.5: Spatial distribution of (a) average monthly NDVI, (b) peak monthly SSEBop ET, and (c) minimum Dwt in June 1999 (wet) and (d) average monthly NDVI, (e) peak monthly SSEBop ET, and (f) minimum Dwt in June 2003 (dry) in the riparian corridor along the Arkansas River.

2.3.3 Relationship between groundwater depth and riparian ET

Values of monthly SSEBop ET averaged over all of the riparian MODFLOW grid cells indicate that May, June, and July are peak months for riparian water use, as shown in Table 2.3. The significant influence of Dwt on riparian ET is highlighted in Figure 2.6, wherein SSEBop ET values computed within individual Landsat cells during June for all eleven years of groundwater modeling are plotted for discrete classes of Dwt values within the respective MODFLOW grid

cells. Based on non-paired Wilcoxon rank sum tests, results show that locations where Dwt is 3 m or less yield statistically significant (α = 0.05) higher ET rates compared to locations with Dwt greater than 3 m. The same significant relationship for Dwt less than 3 m occurs during the months of May and July. The data were tested for normality using the Shapiro-Wilk test [68] to conclude that a non-paired Wilcoxon rank sum test was appropriate.

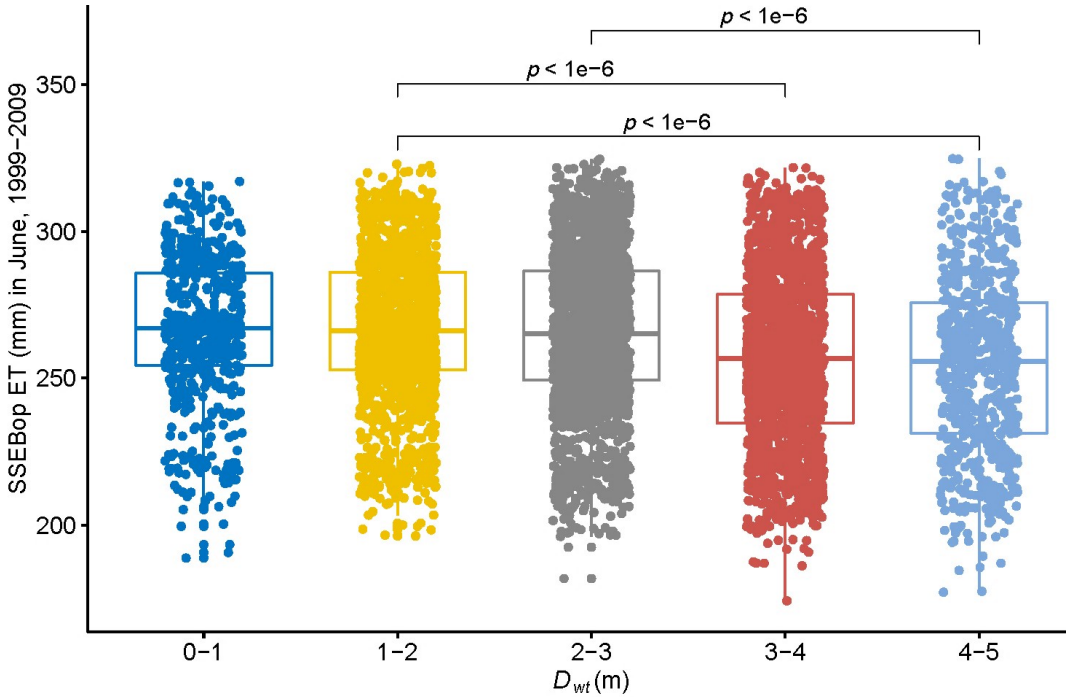

Figure 2.6: Monthly SSEBop ET in June, for all 11 years analysed (1999–2009), divided into 1-m bins of Dwt values. Values of Dwt exceeding 5 m are excluded from this analysis. The p value indicates that a comparison between two group mean values is statistically significant (α = .05).

Values of ρ calculated for pairs of SSEBop ET and Dwt across the grid cells for the months March–October in the wet year 1999 and in the dry year 2003 are plotted in Figure 2.7. Values of ρ outside the range of the shaded zone in Figure 2.7 are statistically significant (α = 0.05) for large sample sizes (n > 100) [69]. The shaded zone represents the region within which ρ values are considered statistically insignificant (600 < n <757, α = 0.05). October 1999 is the only time period during which the value of α is insignificant. Generally, larger negative values

of ρ are present throughout March–October during 2003 compared to 1999, indicating a stronger relationship between high ET values and small Dwt.

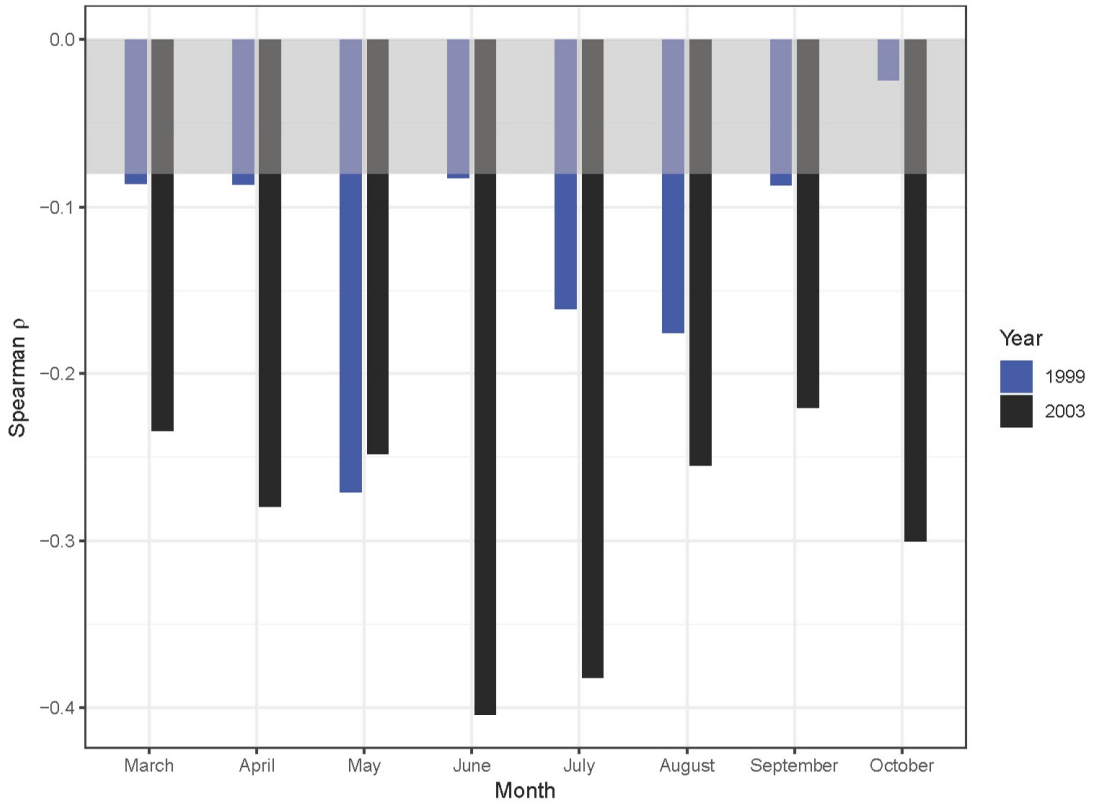

Figure 2.7: Values of ρ for SSEBop ET and Dwt within the riparian corridor for each of the months of the growing season for the wet year 1999 (757 data pairs per month) and the dry year 2003 (600 data pairs per month).

Figure 2.8 depicts frequency histograms of ρ values during 1999 and 2003 within individual MODFLOW grid cells using the eight pairs of SSEBop ET and Dwt values corresponding to the growing season months of March–October. Because only eight pairs of data are compared in the ρ analysis, there are insufficient data to confirm the statistical significance of the results. Generally, ρ values are negative for 1999 and 2003, although more positively-correlated relationships are found in 1999 when compared to 2003.

Figure 2.9A and 2.9B depict the temporal variability of SSEBop ET and average monthly Dwt within the growing season in 1999 and in 2003, respectively. As shown in Figure 2.9, there is more variability in SSEBop ET during the growing season of 2003 compared to 1999. Peak ET rate

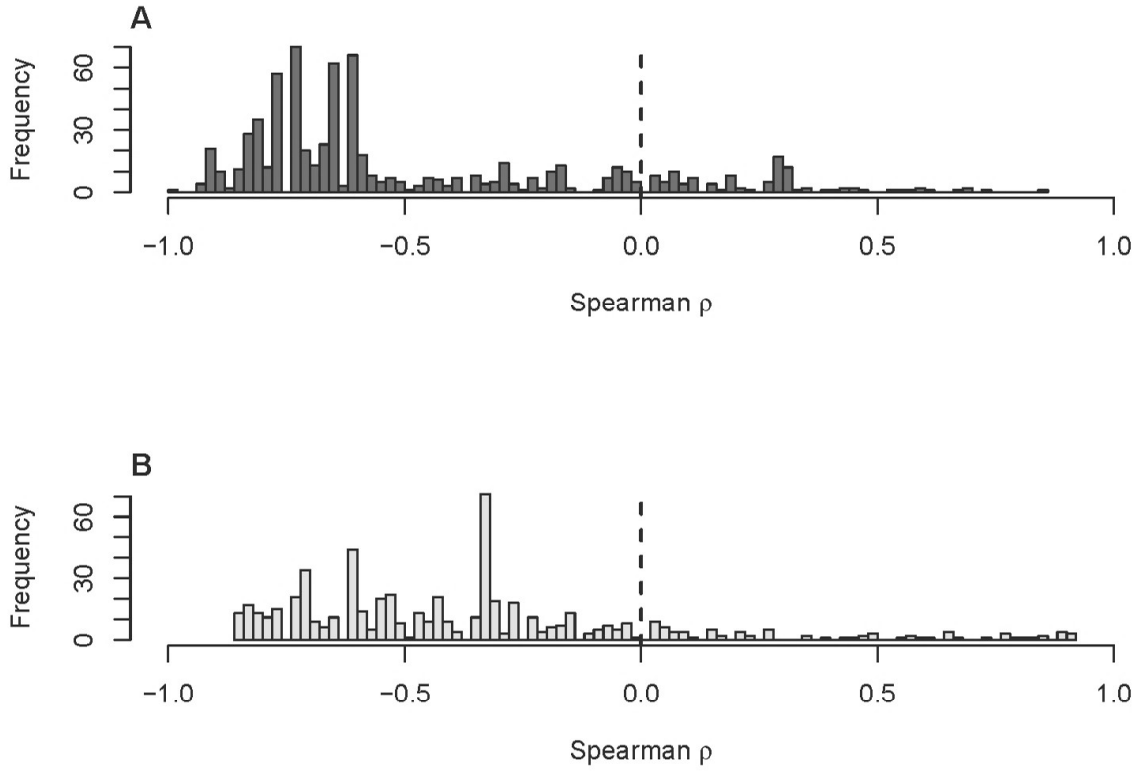

Figure 2.8: Frequency distributions of within-cell ρ values for SSEBop ET and average monthly Dwt, calculated across the eight months of the growing season for (a) the wet year 1999 and (b) the dry year 2003. There are 718 cells with eight data pairs for 1999 and 590 cells with eight data pairs for 2003. There were no tied ranks computed for within-cell ρ values. A vertical line (dashed) is shown on each panel to demarcate the change point between inverse and direct correlations.

occurred in June for 1999 with an arithmetic average of 262 mm, and peak ET occurred in July for 2003 at 298 mm. The shallowest Dwt during 1999 occurred in May with an arithmetic average of 1.9 meters, while June in 2003 showed an average depth of 3.1 meters at its shallowest period, both occurring the month before peak SSEBop ET is observed in each respective year. Groundwater table fluctuations remained shallower in 1999 when compared to data in 2003. Mean-monthly SSEBop ET values were higher in 2003 for six months of the eight-month growing season. Differences between monthly ET rates in 1999 and 2003 are more pronounced during the green-up stages of the vegetation's lifecycle, although agreement between the two growing seasons is observed during annual senescence with differences in mean SSEBop ET as low as 2 mm in October.

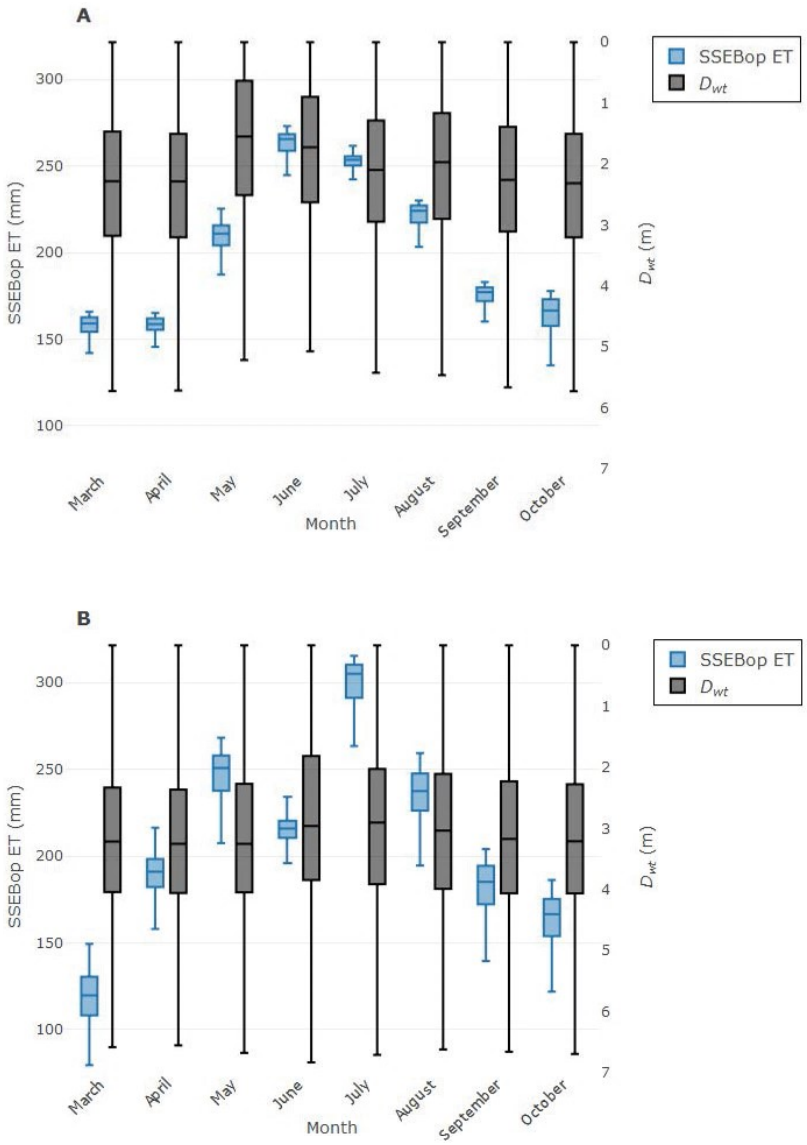

Figure 2.9: Box-and-whisker plots depicting the temporal variability in monthly SSEBop ET with average monthly Dwt for each month within the growing season for (a) 1999 and (b) 2003. The box portion of the box-and-whisker displays 0.25, 0.5, and 0.75 quantile values. The length of the whiskers in the box-and-whisker plot is defined as one and a half the difference of the interquartile range (1.5 [0.75–0.25]). Outliers are not plotted for SSEBop ET and Dwt to focus on values of groundwater depth that are of concern to the dominant vegetation types seen in the study area.

2.4 Discussion

2.4.1 Applicability of SSEBop in riparian areas

Riparian water use continues to be an important area of research as large efforts are directed at controlling invasive riparian species [49, 70] and competing sectors strain for already limited

water availability. Tools that can provide accurate estimates of riparian ET at large spatial scales, such as SSEBop, can help inform water managers of the role of riparian consumptive water use. Senay et al. (2017) applied SSEBop using 31 years of Landsat archives in the southwestern United States and found that riparian and agricultural ET, on monthly and annual timescales, are higher than fallow and non-agricultural landscapes (1800-2000 mm/yr), agreeing with the findings but differing in on monthly estimates, likely due to differing climatic conditions captured by gridMET ETr. Khand et al. (2017) showed the use of a single-satellite-scene (SSS) method, with NDVI as a key input, to have accurate comparisons to a sophisticated model over appropriate spatial scales and subsequently used the SSS method to estimate annual riparian ET for 23 years. Their average annual water use on the Colorado River in southern California is 748 mm which is 62 percent less than the simulated average for all years at 1,949 mm (e.g. Table 2.3). The large difference is likely explained by the fact that vegetation-index methods usually reflect only the transpiration component of ET [71] whereas land surface temperature captures energy fluxes from soil and vegetation.

A promising correlation between daily estimates of SSEBop ET and ETr from a precise-weighing lysimeter (Figure 2.3) supports of the use of the SSEBop methodology in southeastern Colorado. A bias in SSEBop estimates are expected given standard usage of a clear-sky net radiation balance in SSEBop that assumes ET occurs 24 hours a day, which results in overestimation of an independent data set that is based entirely on mass balance. A strong correlation between spatially-averaged riparian SSEBop ET and locally-measured CoAgMET ETr shows that heavily-irrigated crops are similar to energy-limited, groundwater-dependent riparian vegetation (Figure 2.4A). The lower correlation in Figure 2.4B and increased correlation in Figure 2.4A implies that gridMET ETr combined with land-surface temperature can capture local conditions measured by a crop-based micrometeorological station.

2.4.2 Rainfall or energy as drivers of riparian ET

The controlling factors of riparian health can be evaluated by comparing the relationship between rainfall and SSEBop ET. Senay et al. (2017) shows weak relationships between SSEBop ET and rainfall on a monthly and annual timescale for major irrigation districts in California and Arizona. Similar to the findings from this study, a weaker relationship is observed between mean monthly SSEBop ET and rainfall. Given the SSEBop model assumptions, the expectation is that rainfall should create a "colder" land surface temperature and subsequently increase SSEBop ET. The opposite is shown in Figure 2.5 where higher rates of ET in 2003 (dry year), compared to 1999 (wet year), are likely due to increased aridity with no constraint on groundwater supply. Glenn et al. (2015) reported a strong relationship between vegetation-index ET and rainfall over grass and shrubland-dominated study sites. They reported a higher correspondence between ET and the enhanced vegetation index (EVI), R^2=0.85, rather than rainfall, R^2=0.63. There are many differences in this study compared to Glenn et al. (2015), which creates a need to continue experimenting with natural-ecosystem ET using algorithms that incorporate thermographic and normalized-difference reflectance imagery. Smaller correlations between SSEBop ET and rainfall indicates that the riparian vegetation is not limited by moisture, considering the proximity to shallow groundwater, influenced by varying irrigation return flows, but instead follows an optimal path, highlighted by Glenn et al. (2011), where riparian vegetation growth matches other environmental factors like incoming radiation and atmospheric temperature. Monthly rainfall as an explanatory micrometeorological value does not play an increasing role in riparian condition (Figure 2.5) like Ta and CoAgMET ETr (as seen in Figure 2.2 and Table 2.3). This means that the riparian vegetation is not responding predominantly to rainfall as a driver of ET, but rather is varying according to energy limitations.

2.4.3 Groundwater depth influence

Correlations between groundwater depths and peak SSEBop ET in the study agree with Baird and Maddock (2005), who conducted a literature review to confirm plant groups found in wet-

land and riparian ecosystems have ET rates of 2-4 mm/day with a non-linear relationship between groundwater depth and ET flux. This highlights the physiologic sensitivities of riparian species to excess water in energy-limited plant communities (Figure 2.6). Phreatophytic communities located in Wyoming are reported to have rooting depths of 1-3 m with water use ranging from 0 - 3.1 mm/day, and likewise in New Mexico, cottonwood and salt cedar communities reported values of groundwater depth that varied from 0.75 to 1.8 m with ET magnitudes reported as percentages of a water balance [19]. The optimal groundwater depths found in this study, contributing to increased SSEBop ET, agree well with previous ecohydrology studies. A significant "drop-off", observed in Figure 2.6, at groundwater depths below 3 m provides evidence that a threshold exists where groundwater table location influences ET rates regardless of shifts in annual climatic conditions. The estimates of dominant Dwt in semi-arid ecosystems, ranging from 1 to 3 m, agree well with similar research efforts by Zhu et al. (2015). These results suggest the need to consider the impact that improved irrigation management (e.g. increased irrigation efficiency and reduced canal seepage), leading to diminished return flows and larger Dwt, would have on lowering the ET rate within riparian corridors. Such an outcome might prove to be a beneficial side effect that would offset increased water consumption that might otherwise result from a proposed measure to enlarge riparian corridors to promote chemical reduction of pollutants in irrigation return flows [72].

The temporal analysis of peak SSEBop ET indicates riparian vegetation in southeastern Colorado is energy-limited (Figure 2.5). Yue et al. (2016) found that Dwt is a controlling factor in groundwater ET but that seasonal ET is best predicted by atmospheric temperature, agreeing with the findings from this study (Table 2.3). The lack of correlation between a mass balance approach to ET, like the method of White (1932), and local climatological variables in Yue et al. (2016), highlights a need to continue investigating water-use dynamics of natural ecosystems to develop encompassing predictor variables for estimating ET at different scales in space and time. Generally, Figure 2.9 shows less variability in SSEBop ET with shallower Dwt in 1999 compared to 2003. Decreased values of Dwt in 1999 are likely due to increased availability of irrigation water and rainfall amounts. The contrast between Dwt values in 1999 and 2003 verifies the assertion

that shallower water tables on a monthly time scale do not necessarily contribute to increased ET rates when compared to increased arid conditions. Higher rates of SSEBop ET are likely due to increased atmospheric drying with no constraint on groundwater supply in 2003 when compared to estimates from 1999, when rainfall was in excess (e.g. Table 2.3 and Figure 2.5).

2.5 Conclusions

This study assessed the relationship between simulated groundwater depths from a calibrated, finite-difference model and ET estimates from a simplified remotely-sensed ET algorithm within the riparian corridor along a 70-km reach of the Arkansas River, Colorado. A key finding is that simulated riparian ET is primarily controlled by local micrometeorological variables associated with energy-availability rather than rainfall. In addition, results revealed that groundwater depths of 3 m or less had significantly higher ET rates during peak growing season months for all 11 years of analysis.

The Spearman correlation between groundwater depth and SSEBop ET during a wet and dry year, on a monthly basis, revealed stronger relationships during drier conditions. The results of this unique, data-rich, dual-model comparison agreed with previous ecological research showing a non-linear relationship between ET and groundwater depth.

Most of the analysis presented here focused on temporal variations between SSEBop ET and groundwater depth, defined by differing hydroclimatic years, with the use of a highly-calibrated groundwater model. To further the understanding on the linkages between floodplain activity and a riparian ecosystem, the next approach includes a statistically-inclined method to examine factors such as spatial proximity to groundwater pumping wells, tributaries, and irrigation canals in relation to natural-ecosystem response. The controlling variables of riparian health change with scale in both space and time, and therefore future efforts will continue to experiment on the multi-scale predictors and responses of groundwater-dependent ecosystems [28].

Chapter 3

Space and Time Riparian Vegetation Characteristics

3.1 Introduction

Many rivers run through vast agricultural regions, often irrigated. Vegetation in the riparian corridor is important for habitat biodiversity [33], removing agricultural-borne pollutants [55], attenuating floods [73], and riparian ecosystems rely on groundwater return flow [28, 74]. Understanding the interactions between an agro-ecosystem and a riparian ecosystem can improve water management that include compliance with standards for water quality, conjunctive use of groundwater and surface water, and conservation of a vulnerable ecosystem.

The knowledge-base on the quantifying the connectivity between an irrigated landscape and riparian water use is incomplete. There is little data about the temporal impact of irrigation-perturbed return flow on riparian water use. Temporal lags and spatial dependencies between the cause (i.e., return flows) and effect (i.e., riparian water use) should be examined to prevent the loss of a vulnerable ecosystem to increased water demand, changing climate, and evolving irrigation methodologies. In southeastern (SE) Colorado, USA floodplain activity consists of flood-furrow irrigation of agricultural crops with a portion of land that utilizes center-pivot and drip technology [52]. Flooding as an irrigation method, as opposed to center-pivot, produces increased return flow. A review of groundwater recharge measurement methodologies by Scanlon, Healy and Cook (2002) shows irrigating activity can contribute a large portion of water to return flow [23]. In turn, locations along the riparian corridor that filter the return flow can depend on the extra flow as a water source [22]. Relevant research estimates 50% of the water applied to agricultural areas in the Arkansas River basin is converted to return flow [57, 75]; and irrigation-altered return flows are connected to riparian ecology along the Arkansas River [59]. Recently, a global synthesis of recharge data indicates that groundwater plays a larger role than previously thought in controlling

streamflow and evapotranspiration [76]; and irrigation communities can contribute a large portion of water to return flow [23].

Two unique layers of complexity to the water balance in SE Colorado consist of a process known as canal augmentation and the influx of non-native vegetation. Canal augmentation is utilized to replace water in timing and amount to a surface water source when the water user requires water and supplies are not sufficient to meet the demand [25]. The augmentation water prevents changes in historic stream hydrology when irrigation methods change. Additionally, unique to river systems in the western USA, non-native species like salt cedar (Tamarix ramnossima) grow unfettered given their ability to withstand high temperatures and salt-laden soils [59, 77, 78]. The research presented here examines the interrelationships between irrigation-influenced hydrologic factors and vegetation patterns along the Arkansas River in SE Colorado.

Publicly available data, Google Earth Engine's cloud computing capabilities, and Bayesian techniques are applied to examine the spatial and temporal dependencies of riparian vegetation health. The models applied in this study deal with lagged covariates and autoregressive (AR) signatures and are widely used in economic, hydrologic, and ecological studies to understand data and make predictions [79–83]. Lagged temporal regression (e.g., estimation of dependence on factors at earlier times) is appropriate for examining riparian water use because the state of the riparian ecosystem at a given time could be due to past inputs into the watershed [22, 84]. Spatial regression combined with categorical covariates can answer questions about the connections between floodplain activity and riparian function. A family of Bayesian linear regression models are used to investigate patterns of vegetation over time and space. Linear models are used because: advanced applications of remote sensing of evapotranspiration (ET) have linear functional forms, and though the underlying mechanisms that connect human-made and natural ecosystems are complex, applying linear models is a necessary step in understanding spatio-temporoal processes.

The central aim in this research is to evaluate what temporal and spatial statistical models can uncover about the connections between natural ecosystem water use and the adjacent landscape and if the findings agree with previous studies investigating this science question. The following

hypotheses are addressed: (1) temporal trend in monthly normalized-difference vegetation index (NDVI) covering the riparian vegetation along the Arkansas River is only detectable at spatial scales smaller than catchment (10-km), (2) the time lag of monthly NDVI with precipitation is shorter than that with river discharge, (3) riparian areas along the Arkansas River that intersect with perennial streams have increased cumulative vegetation density compared to areas abutted by permanent land-use change (e.g., agriculture transition to fallow). These hypotheses are proposed to deduce the hydrosocial connections between human-built and natural ecosystems to inform decision making at operational scales. The analysis conducted herein tests these predictions using a time series modeling approach, previously used to highlight the connections between vegetation life cycles and various external forcings [21, 24, 80, 85–87], and a spatial modeling approach with Gaussian-process regression techniques that has been previously used on spatial data problems to account for spatial random effects between similar regions [88–90].

3.2 Methods

3.2.1 Study Area

The riparian vegetation under examination borders a 96-km reach of the Arkansas River in SE Colorado with an areal coverage of 52.5 km^2. Figure 3.1 depicts the study area that begins downstream of the John Martin Reservoir near Lamar, Colorado and extends to the Colorado-Kansas border. The study area is located in a semiarid climate with rainfall ranging from 343 mm to 394 mm [25]. The shallow alluvial aquifer, in connection with the Arkansas River, supports the riparian vegetation as a groundwater-dependent ecosystem. The natural-vegetation extent of the riparian corridor is demarcated using Google's platform of satellite imagery (varies upon resolution) and a series of past cropping-pattern extents from the National Land Cover Database collected in 2001 and 2016.

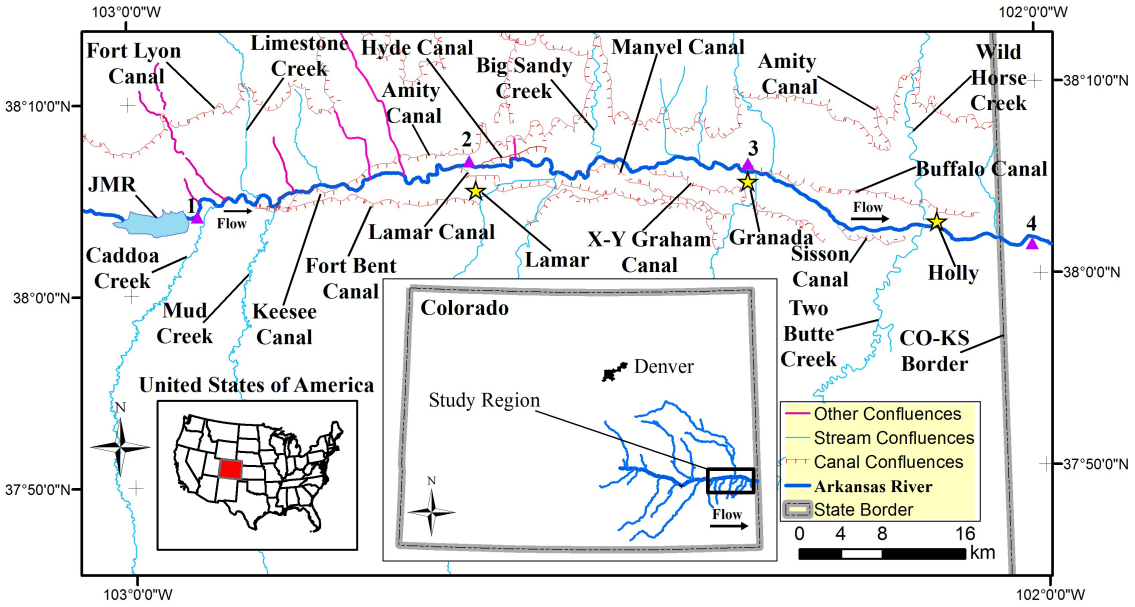

Figure 3.1: General location map of the 96-km reach of the Arkansas River located downstream of John Martin Reservoir (JMR). Yellow stars indicate a township or city and the purple triangles represent four U.S. Geological Survey (USGS) gages used in this study.

3.2.2 The Landsat Data and Strategic Spatial Partitioning

Using Google Earth Engine (GEE), all Landsat 5, 7, and 8 imagery (path: 31 and row: 34) from 1984-2019 are used to extract NDVI for the study area. The NDVI product is used in this study to track natural-ecosystem biomass, as a proxy for evapotranspiration, due to the history the product has as an effective parameter in satellite-based evapotranspiration algorithms [38, 47] and because NDVI has been validated in tracking groundwater-dependent ecosystems [20, 28]. In total, data from 932 satellite images from GEE are used. Each satellite pixel (30 m) centroid that falls within the areal extent of the riparian corridor is defined as a sampling point to extract values of surface-reflected NDVI from each satellite scene. This sampling scheme resulted in $58,443$ total satellite pixels covering the study area. All imagery is screened for clouds, ice and cloud shadows within the areal extent of the riparian corridor following the same methodology as Albano et al. (2020) [12]. Following Hungtington et al. (2016), Landsat 8 Operation Land Imager data is linearly corrected for comparison with Landsat 5 and 7 [91]. Similar corrections have been applied to Landsat imagery covering riparian vegetation in Vanderhoof et al. (2019) and Albano et al. (2020). The data are exported from GEE into R to be organized in chronological order (R

Core Team, 2018). All pixel values found in error, likely due to the presence of clouds or snow, are assigned an NA value. Points that fall within the water surface area of the Arkansas River are removed using the NLCD Open Water classification. All sampling points are divided into areal plots, hereafter referred to as *subregions*, based on the location of the sampling point along the Arkansas River.

Average NDVI for each month of the year at each Landsat pixel is computed using the *xts* package in R to prepare data for the independent space and time analyses [92]. Pixels that contain an NA value in the time series are excluded from the calculation of the monthly average. Temporally averaging all data to a value per month compensates for gaps in satellite imagery from clouds and snow and reduces noise [24, 93]. The study area is split into subregions using demarcations that correspond to areas with eight distinct hydrologic response functions as defined by the rainfall-runoff model that Colorado's Division of Water Resources (DWR) uses to model sections of the Arkansas River with the Hydrologic Institutional Model to ensure compliance with the Arkansas River Compact between Colorado and Kansas [94]. Table 3.1 provides a summary of the subregions within the study area.

Table 3.1: A summary of the subregion demarcations that correspond to the Hydrologic Institutional Model to simulate water processes on the Arkansas River. The Δ Dist. (km) is the incremental length of each river segment in a subregion.

No.	Name	Location (Long.°)	Δ Dist. (km)
1	Below JMR to Below Fort Bent Canal	(begin,-102.8°]	10.03
2	Below Fort Bent Canal to Below Amity Canal	(-102.8°,-102.7°]	11.38
3	Below Amity Canal to Lamar Gage	(-102.7°,-102.6°]	10.48
4	Lamar Gage to Below Manvel Canal	(-102.6°,-102.5°]	17.39
5	Below Manvel Canal to Below X-Y Graham Canal	(-102.5°,-102.4°]	8.78
6	Below X-Y Graham Canal to Granada Gage	(-102.4°,-102.3°]	11.28
7	Granada Gage to Below Sisson-Stubbs Canal	(-102.3°,-102.2°]	14.43
8	Below Sisson-Stubbs Canal to CO-KS Border	(-102.2°, end]	12.80

Temporal Representation of NDVI

Each subregion is represented as a time series vector with dimensions of $T = 428 \times 1$, where T is the length of the time series in months. Each value in the vector is average monthly NDVI ranging from May 1984 to December 2019. Subregions ($d = 1, ..., D = 8$) represent the number of separate time series vectors created for the entire study area. Monthly NDVI is computed at the Landsat-pixel level first, and then the areal average of all satellite pixels located within a given subregion creates a single average NDVI value per month per subregion.

Spatial Representation of NDVI

Temporal integration (i.e., sum) is applied to average monthly NDVI to convert the time series into a single value of cumulative density at every Landsat pixel within the study region. Previous research suggests that integrated NDVI is strongly correlated to aboveground net primary production in arid regions [95]. The approach to integrated NDVI included herein is simpler than the methods proposed by Reed, Loveland and Tiezen (1996), who defined the onset of greenness and summed all values of NDVI to a point in time when the plant is considered dormant [96]. The temporal integration is applied to all available months of data (2000-2019) to utilize all data and to limit user selection bias of seasonal start and stop points.

The time period chosen to analyze temporally-integrated NDVI is from January 2000 to December 2019. By 1996 several major surface-irrigation canals in the study area had been affected by augmentation [25], so January 2000 serves as an inflection point to test if integrated NDVI will show distinct patterns in proximity to augmented canals along the Arkansas River. Satellite pixels are filtered in space and time to normalize each subregion and thereby create an equal-area analysis among subregions (i.e., all subregions have the same number of observations of integrated NDVI per satellite pixel). The data filters are as follows: exclude data in space by removing all pixels with five or more observations of monthly NDVI less than 0, and exclude data in time by removing images with greater than 1% of missing areal data within the riparian delineation. The filtered data are randomly sampled to create a data vector for the spatial analysis that is of size $N_d = 500 \times 1$ (0.45 km^2) where N_d is the number of spatial observations in a given subregion.

3.2.3 Temporal Covariates

Cumulative monthly precipitation, and average daily temperature per month, for the Upper Arkansas hydrologic boundary (HUC: 11020002) are downloaded from WestWide Drought Tracker [97]. Based on natural ecosystem ET models and a long record of recordings in the study area, atmospheric temperature is used to account for the seasonality in the NDVI data [17]. Precipitation can play an early and direct role in riparian ecosystem ET [98]. Precipitation and temperature are assumed spatially uniform in the study area (i.e., all subregions share the same data). Temperature (T) and precipitation (P) are standardized before being analyzed to reduce collinearity and improve model convergence [89].

Monthly mean discharge for the Arkansas River is extracted from USGS Surface-Water Monthly Statistics for the Nation [99]. To reduce collinearity between precipitation and river discharge (Q), four USGS gages identified in Figure 3.1 are used to perform a mass balance from the upstream gage to the nearest downstream gage. This mass-balance technique converts discharge into a value of net stream gain or stream loss per month (Q_{G/L_t}), hereafter referred to as stream gain-loss. For example, if the upstream gage is recording a flow that is less than the nearest downstream gage, then that reach segment is gaining by the difference between the two flows. The four USGS gages identified by USGS ID are: Below JMR (ID: 07130500; 38.07°, -102.93°), at Lamar, CO (ID: 07133000; 38.11°, -102.62°), near Granada, CO (ID: 07134180; 38.10°, -102.31°), near Coolidge, KS (ID: 07137500; 38.03°, -102.01°). The stream gain-loss covariate is used as the key proxy to connect the agroecosystem to the riparian ecosystem. The assumption is that when the river is gaining then return flow (i.e., discharge from the alluvium to the river) is occurring.

The portion of the Arkansas River from the outlet of the John Martin Reservoir to the CO-KS border is highly managed. Given that the John Martin Reservoir is the only on-stream reservoir in the study area, the reservoir is considered a hydraulic boundary. Reservoir storage data is used as a temporal covariate because of the close connections reservoir storage has to river discharge and applied irrigation water [25]. Reservoir storage data for John Martin Reservoir are collected from the geographical information system and mapping tools hosted by Colorado DWR. The daily

storage in acre-ft is averaged to monthly storage and converted to cubic meters. The storage data are standardized before being used as input to the time series model.

Data on monthly depth to groundwater (*dtgw*) are also collected from the geographical information system hosted by DWR. There are numerous groundwater wells in the study area. One groundwater observation well is chosen for each subregion based on the following criteria: duration of observations is from 1980-2018, the observation well is within the spatial extent of the Quartenary Alluvium, shortest distance to the Arkansas River, and within the longitudinal extents of a subregion (see Table A.1 in the Appendix). The only exception to this criteria is the groundwater observation well used for subregion 8 which is located at the downstream end of subregion 7. This inconsistency is because subregion 8 did not have observation wells that met the criteria above. Each observation record had to be extended to 2019 so the WestWide Drought Tracker [97] is used to find a similar hydrologic year, based on percentage of normal for precipitation, and the groundwater data from the hydrologically-similar year replaced the missing data in 2019. The Flood Emergency Management Agency's Flood Map Service Center determined if each groundwater observation well is within the floodplain. Four monitoring wells corresponding to subregion 3, 4, 5 and 7 are located within a flood risk area (Zone A, AE, X), two monitoring wells corresponding to subregion 1 and 2 are located in a non-digitally mapped area, and two monitoring wells corresponding to subregion 6 and 8 are located on the edge of Zone A flood risk extents.

3.2.4 Spatial Covariates

Using CDSS's Map Viewer, the spatial location of major confluences with the Arkansas River are collected for the analysis (see Table A.2 in the Appendix). Each explanatory variable used in the spatial analysis is called a group and each subregion is assigned a value from each group. Data for several variables are compiled to explain spatial changes in temporally-integrated riparian NDVI per subregion. The groups are described in order as they appear in Table 3.2: (1) number of confluences, (2) are the operating canals outside of the riparian corridor impacted by augmentation [*yes* (1) or *no* (0)], (3) is there development (i.e., city or pastoral practices) located outside of the

adjacent floodplain [*yes* (1) or *no* (0)], (4) is there a confluence between the Arkansas River and a perennial tributary [*yes* (1) or *no* (0)], and (5) is the subregion adjoined by permanent dry up locations of irrigated land identified by the Substitute Water Supply Plans or decrees in Colorado for the Division of Water Resources [*yes* (1) or *no* (0)] (see Table A.3 in the Appendix for canals impacted by augmentation).

The spatial analysis includes count and binary variables that can be used to verify predictions about the connections between a riparian ecosystem and the surrounding landscape (Table 3.2). The number of confluences is used as a count-covariate to account for the presence of drainage ditches, canals and ephemeral draws in a subregion as an indication for seasonally-shallow groundwater (Group 1). Canal augmentation is coded as a binary covariate to track changing water availability in a subregion as a result of altered soil-moisture and infiltration from modified irrigation methodologies (Group 2) [24]. Floodplain activity is a broad covariate that tests if pastoral practices and cities impact temporally integrated NDVI through altered surface runoff in timing and amount (Group 3). Confluence type is used as a binary covariate to discover the impact of increased water availability from perennial confluences on cumulative vegetation density (Group 4). Irrigated land dry-up indicate subregions that are abutted by historic agricultural land that are no longer irrigated but can be re-irrigated with groundwater under a plan for augmentation, dryland farmed or have revegetation requirements (Group 5).

Table 3.2: A summary of the subregions with count and binary variables incorporated into a coding matrix.

Subregion	Group 1	Group 2	Group 3	Group 4	Group 5
1	4	0	0	0	0
2	3	1	0	1	1
3	2	1	1	0	0
4	2	0	1	0	0
5	1	1	0	1	1
6	2	1	0	0	0
7	2	0	0	0	1
8	3	1	0	1	0

3.2.5 The Linear Models

The Bayesian multiple linear regression models used in the analyses are shown in Table 3.3 below. Model 1 is a classic linear regression model in a Bayesian setting. Model 2 is a hierarchical Bayesian linear time series model that accounts for an autoregressive signature of order one. Model 3 is a hierarchical Bayesian linear network model that accounts for spatial dependence among pixel neighbors within a given subregion. The reader is directed to Hooten and Hefley (2019) for a more in-depth understanding of the models used in this study. A custom Markov chain Monte Carlo (MCMC) sampler is created for the models shown in Table 3.3 following the six general steps outlined by Hobbs and Hooten (2015). Appendix A shows the model posterior distribution for all three models (Section A.1) and the full conditional statements for all model variables (Section A.2), respectively. Model 1 has an update for the vector of correlation coefficients (β_i) and the uncorrelated error variance (σ^2). Model 2 and 3 deal with correlated and uncorrelated variances in the temporal and spatial datasets.

Table 3.3: A summary of the models used in this study. The symbol $\sim$ implies two sides of an equation are proportional to each other. The probability distribution abbreviations are as follows: Normal-Gaussian (N), Multivariate Normal (MVN), and Inverse Gamma (IG).

Model 1 (General)	Model 2 (Time)	Model 3 (Space)
$\mathrm{y_i} \sim \mathrm{N}(\mathbf{x}_i^{'}\boldsymbol{\beta}, \sigma^2)$	$\mathrm{y}_t \sim \mathrm{N}(\mathbf{x}_t^{'}\boldsymbol{\beta} + \mathrm{z}_t, \sigma^2_{\mathrm{y}})$	$\mathbf{y}_d \sim \mathrm{MVN}(\mathbf{X}_d'\boldsymbol{\beta} + \boldsymbol{\eta_d}, \sigma^2_d\mathbf{I})$
$\boldsymbol{\beta} \sim \mathrm{MVN}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$	$\mathrm{z}_t \sim \mathrm{N}(\alpha \mathrm{z}_{t-1}, \sigma^2_{\mathrm{z}})$	$\boldsymbol{\eta_d} \sim \mathrm{MVN}(\boldsymbol{\mu}_\eta, \boldsymbol{\Sigma}(\sigma^2_{\eta_d}, \rho))$
$\sigma^2 \sim \mathrm{IG}(q, r)$	$\boldsymbol{\beta} \sim \mathrm{MVN}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$	$\boldsymbol{\beta} \sim \mathrm{MVN}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$
	$\alpha \sim \mathrm{N}(\mu_\alpha, \sigma^2_\alpha)$	$\rho \sim \mathrm{Gamma}(\gamma_1, \gamma_2)$
	$\sigma^2_{\mathrm{y}} \sim \mathrm{IG}(q_{\mathrm{y}}, r_{\mathrm{y}})$	$\sigma^2_d \sim \mathrm{IG}(q_d, r_d)$
	$\sigma^2_{\mathrm{z}} \sim \mathrm{IG}(q_{\mathrm{z}}, r_{\mathrm{z}})$	$\sigma^2_{\eta_d} \sim \mathrm{IG}(q_\eta, r_\eta)$
	$\mathrm{z}_0 \sim \mathrm{N}(\mu_{\mathrm{z}_0}, \sigma^2_{\mathrm{z}_0})$	

Model 2 equates to a fully-Gibbs MCMC algorithm (Appendix A2 and B2). The model variables for the time series case are described as follows: y_t comes from a univariate normal distribution and represents a measurement of monthly NDVI at time $t = 1, ..., T$, $\mathbf{x}_t$ is a $p \times 1$ vector of fixed covariates measured at time t where p represents the number of covariates, $\boldsymbol{\beta}$ is a multivari-

ate normal $p \times 1$ vector of regression coefficients with a $p \times 1$ prior mean vector $\boldsymbol{\mu}_\beta$ and a $p \times p$ covariance matrix $\boldsymbol{\Sigma}_\beta = \sigma_\beta^2 \mathbf{I}$, each $\mathbf{I}$ is an identity matrix with 1 on the diagonal elements of the matrix and 0 elsewhere, z_t represents the autoregressive error in the ecohydrologic process at time t, and α is a scalar autocorrelation parameter of order one with hyperparameters μ_α and σ_α^2. The data model variance, σ_y^2, accounts for the leftover variance not captured by the mean and autocorrelation components. The process model variance, σ_z^2, is a scalar that represents the variance in the autoregressive error, and z_0 is the initial state which comes from a univariate normal distribution. A directed acyclic graph is shown in Figure 3.2 as a visual guide for Model 2.

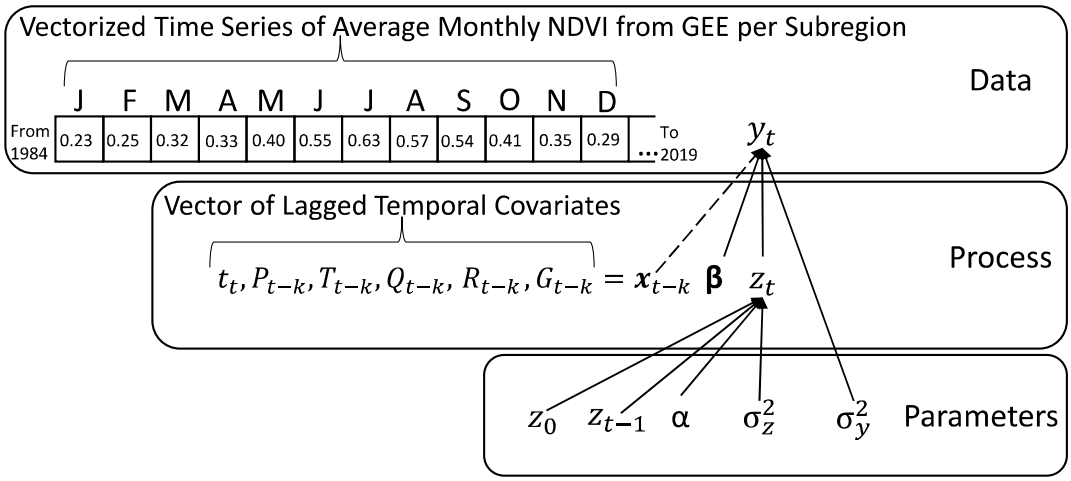

Figure 3.2: A directed acyclic graph for the model of the dynamics of riparian NDVI along the Arkansas River. The solid lines represent stochastic relationships and the dashed lines represent deterministic relationships.

The spatial model is a hybrid Gibbs and Metropolis-Hastings sampling scheme (Appendix A3 and B3). The spatial model is similar to the time series model in that the model accounts for a latent correlated process directly (i.e., z_t in the temporal model and $\boldsymbol{\eta}_d$ in the spatial model). The correlated process ($\boldsymbol{\eta}_d$) comes from a multivariate normal distribution; $\boldsymbol{\mu}_{\eta_d}$ is the $n_d \times 1$ prior mean vector and $\boldsymbol{\Sigma}(\sigma_{\eta_d}^2, \rho)$ is an $n_d \times n_d$ covariance matrix set equal to $\sigma_{\eta_d}^2 \mathbf{R}(\rho)$ which is equivalent to $\sigma_\eta^2(\text{diag}(\mathbf{W}\mathbf{1}_d) - \rho\mathbf{W}_d)^{-1}$; $\text{diag}(\mathbf{W}\mathbf{1}_d)$ is a diagonal matrix with the row sums of $\mathbf{W}$ as diagonal elements [79]. This network model has $\mathbf{W}_d$ as a binary distance matrix where locations

within a certain distance of a pixel are neighbors [79]. In the neighborhood matrix, locations that are neighbors are assigned a one, and a zero otherwise. In this study the binary proximity matrix is developed for each subregion by dictating that all pixels within 45 m of each other are considered neighbors (e.g., similar plant functional group); and Moran's I confirms positive spatial structure in the data, per subregion, using the *ape* package [100, 101] at grid cell sizes greater than 30 meters. Correspondingly, literature indicates Canadian riparian forests can be accurately classified at spatial resolutions of 8.4 m and 30.8 m [102]. The spatial dependence parameter, ρ, is from a symmetric proposal with a univariate normal distribution set as the mean and a variance with a tuning parameter [79]. In the spatial formulation, σ_d^2 represents the left over variability after accounting for correlation among pixels that represent the same plant functional group. The variance in the spatially correlated random effect, $\sigma_{\eta_d}^2$, is the expected error between the predictions of temporally integrated NDVI for image pixels that are neighbors and contain a similar plant functional group. The regression coefficients ($\boldsymbol{\beta}$) and the spatial dependence parameter (ρ) are shared among subregions which forces smoothing across the eight subregions [90]. A directed acyclic graph is shown in Figure 3.3 as a visual guide for Model 3.

The Priors

The prior distributions represent insight about a variable in a statistical model and does not depend on the data being analyzed. All variance structures are specified with strictly positive distributions (e.g., Gamma and IG). The autocorrelation coefficient in the time series model (α) is specified with a mean (μ_α) equal to 0.6 and a variance (σ_α^2) equal to one because the correlation parameter is between -1 and 1 for stationary time series. The hyperparameters used to explain the covariance structures are vaguely specified with $q_{\mathrm{z}} = 0.001$, $r_{\mathrm{z}} = 1000$, $q_{\mathrm{y}} = 0.001$ and $r_{\mathrm{y}} = 1000$. The initial state variable has a mean of $\mu_{z_0} = 0$ and a variance $\sigma_{z_0}^2 = 1$. The mean vector for the regression coefficients $\boldsymbol{\mu}_\beta$ is set equal to zero and $\boldsymbol{\Sigma}_\beta$ is equal to $10000\mathbf{I}$ to control collinearity among covariates [79].

The priors in the spatial model are weakly informed similar to the time series approach, except for the spatial dependence parameter (ρ) and the spatially-correlated process ($\boldsymbol{\eta}_d$). The spatial

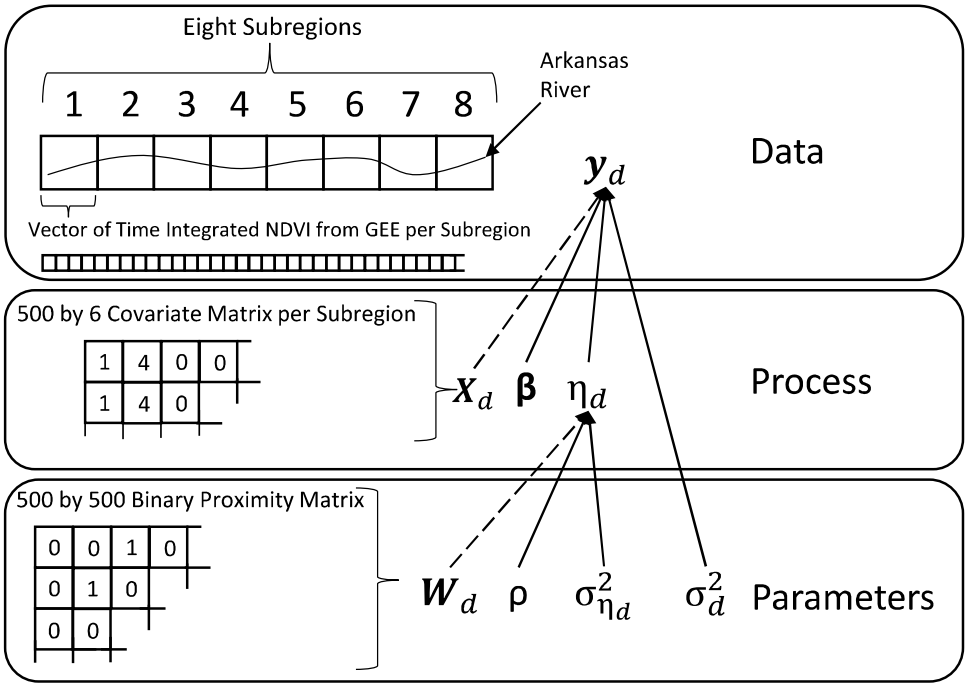

Figure 3.3: A directed acyclic graph for the spatial model of time integrated riparian NDVI along the Arkansas River. The solid lines represent stochastic relationships and the dashed lines represent deterministic relationships.

dependence parameter has hyperparameters set to $\gamma_1 = 9801$ and $\gamma_2 = 9900$ and ρ is instantiated in the MCMC chain very close to one (e.g., 0.999) with a very small tuning parameter ($\rho_{\text{tune}} = 0.0005$). The MCMC algorithm has an accept-reject sampling scheme so that the spatial dependence parameter remains between 0 and 1 [79]. The model is informed so that $\sigma^2_{\eta_d}$ comes from a mean of 5 and a variance of 4. This informed prior helps the model identify the spatial dependence among like vegetation communities as the smaller covariance between two covariance structures (e.g., σ^2_d and $\sigma^2_{\eta_d}$). The independent error variance, σ^2_d, is vaguely informed with $q_{\text{d}} = 0.001$ and $r_{\text{d}} = 1000$.

Convergence, Model Checking and Model Comparison

All models used in this analysis are checked for convergence by visual inspection of the MCMC trace plots. Posterior predictive distributions are used to determine if the models are capable of generating data that resemble the observed data [89]. The deviance information criterion (DIC) is used to measure how each covariate impacts the model in the time series approach. The DIC score

is commonly used in Bayesian analysis and is good for comparing models of the same class [89]. The DIC is calculated as:

$$\text{DIC} = -2\log[\mathbf{y}|\mathbf{E}(\boldsymbol{\theta}|\mathbf{y})] + 2p_D \tag{3.1}$$

where the term -2log[$\mathbf{y}|\mathbf{E}(\boldsymbol{\theta}|\mathbf{y})$] represents the loss function, and $2p_D$ represents a penalty based on number of model parameters [79]. The model with the lowest DIC score fits the data the best [79].

3.2.6 Prediction Testing

Hypothesis 1: Identifiable Temporal Trend At Catchment Scale

To determine the spatial scale at which temporal trend is identifiable, the study area is divided along the Arkansas River into successively smaller areas and Model 2 is used to derive the correlation coefficient for the trend covariate. The *baseline* version of Model 2 consists of an intercept, trend, periodicity, an AR(1) component, and an independent error component (e.g., $y_t = \alpha z_t + \beta_0 + \beta_1(t) + \beta_2(T_t) + \epsilon_t$). The study adopts definitions of scale in hydrology from Blöschl and Sivapalan (1995) [30]. The study area is divided into sub-areas with resolutions of: large-reach scale (1 km), catchment scale ($>$ 10 km), and sub-regional scale (100 km). If the 95% credible interval of the posterior distribution for β_1 does not overlap zero, this suggests there is growth or decay in time at a given spatial resolution. I predict that temporal growth or decay is identifiable at scales smaller than catchment because of first-hand experience with the study area and previous research on riparian trend [12, 85].

Hypothesis 2: Precipitation Will Have A Shorter Lag Effect

The DIC score (Equation 3.1) for each subregion is used to authenticate which hydrologic covariate has the shortest lag impact on monthly NDVI. This work focuses on the intra-seasonal impacts of hydrosocial covariates on riparian NDVI. Pearson correlation between covariates confirmed obvious correlations among precipitation and stream gain-loss. Therefore each covariate, as a set of intra-season lags (i.e., t, $t-1$, and $t-2$), is added to the baseline model separately to determine the predictive ability of each covariate. There are seven possible combinations of a

lagged covariate using times t, $t-1$ and $t-2$. The DIC score helps remove lagged covariates that cannot explain variability in the data. Precipitation is assumed to have a shorter lag effect because recent literature, studying a riparian ecosystem with similar climatic conditions, suggest that there are stronger relationships between riparian NDVI and precipitation compared to other hydrologic inputs like streamflow and reservoir stage [12].

Hypothesis 3: Confluences With Perennial Streams Increase Cumulative Vegetation Density

A spatial random effects model is used with a spatial partitioning approach, similar to the one-of-many applied in Heaton et al. (2019), to test how large scale floodplain characteristics impact riparian vegetation [90]. The perennial confluence descriptor variable serve as a hydraulic boundary that uplifts the local groundwater table creating a return flow hotspot that supports existing and emerging vegetation. If the posterior distribution of the model parameter significantly overlaps zero then the influence between that covariate and integrated NDVI is undetectable, given the other covariates in the model.

3.3 Results

For the temporal analysis, applicable to the first and second hypothesis, the satellite-pixel averaging that creates a monthly time series results in a range of 157 to 232 missing values from a given subregion's time series. After analyzing each time series, derived distributions for model parameters are created using MCMC. Previous examples of the derived distribution approach in hydrologic science and water-resources engineering are Eagleson (1978), Ramirez and Senarath (2000), and Chen et al. (2008) [103–105]. I present the time series and spatial statistics results in a similar fashion.

3.3.1 Hypothesis 1: Identifiable Temporal Trend At Catchment Scale

To address the first prediction, the study area is split into progressively finer scale riparian areas to determine the spatial-scale threshold for detecting trend. Figure 3.4 shows the posterior distribution for the trend covariate under varying spatial resolutions. Temporal trend, positive or

negative, is not detectable at spatial scales larger than catchment size ($>$ 10 km). At spatial scales less than catchment size, 20% of sub-areas analyzed show identifiable trend, growth or decay, from 1984 to 2019.

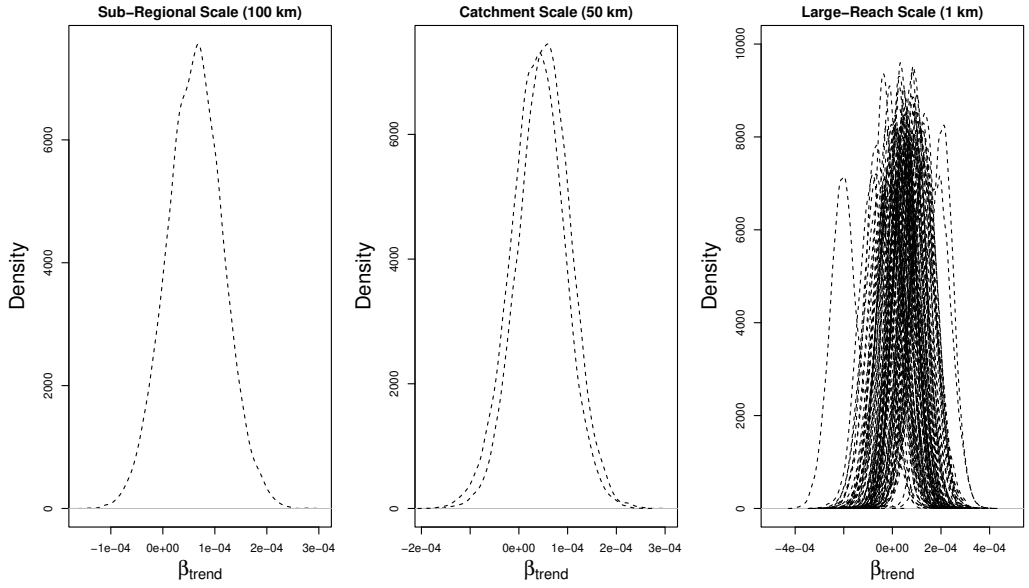

Figure 3.4: Marginal posterior densities for β_{trend} estimated at varying spatial resolutions.

3.3.2 Hypothesis 2: Precipitation Will Have A Shorter Lag Effect

The DIC scores in Table 3.4 partially corroborate the second prediction by indicating that the impact of precipitation on riparian vegetation has a shorter lag time than stream gain-loss but is dependent on subregion. Most subregions share correlation with intra-seasonal precipitation and stream gain-loss at similar time lags but subregion 6 is the only area studied that exhibited a shorter lag between NDVI and precipitation. In the simulation that tested the simultaneous predictive ability of P at t, $t-1$, and $t-2$ at each subregion, precipitation showed direct correlations with monthly NDVI regardless of subregion. Subregions 1, 3 and 8 are positively correlated with P at time t, $t-1$, and $t-2$. Subregion 4, 5 and 7 are the only subregions that had more predictive DIC scores using the most parsimonious model (i.e., intercept and seasonal components). The posterior

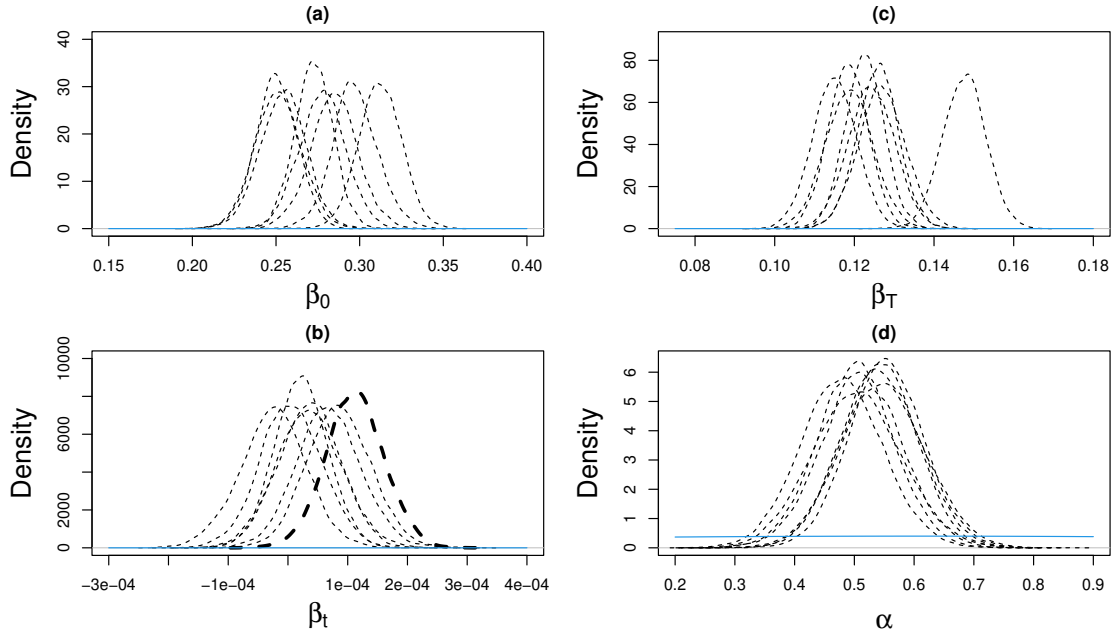

Figure 3.5: Marginal posterior densities for β_0, β_{trend}, β_{Temp}, and α. There are eight histograms per panel. The prior mean ($\boldsymbol{\mu}_\beta = 0$, $\boldsymbol{\mu}_\alpha = 0$) is shown as a blue line. In panel B, the central 95% Bayesian credible intervals do not overlap zero for subregion three (below Amity Canal to Lamar Gage) which indicates a positive trend. Subregion three in panel B of the figure is highlighted using a bold-dashed line.

distributions (not shown) for precipitation show that the strongest positive correlation between the response variable and precipitation is at time $t-1$.

Monthly NDVI are correlated with stream gain-loss but it depends on subregion. In the simulation that tested the simultaneous predictive ability of $Q_{G/L}$ at t, $t-1$, and $t-2$ at each subregion, stream gain-loss showed direct and indirect correlations with monthly NDVI depending on subregion. Subregions 1, 2, 3, and 8 are correlated with $Q_{G/L}$ at time t. Subregions 1, 2, 3, 4 and 6 are correlated with $Q_{G/L}$ at time $t-1$. Subregions 1, 4, and 6 are correlated with $Q_{G/L}$ at time $t-2$. Subregions 5 and 7 showed no correlation with any lags of stream gain-loss (Table 3.4).

Other covariates are examined to determine the temporal relationship with the response variable. The posterior distribution for trend correlation coefficient (β_t) overlaps zero for all subregions except for subregion 3 (below Amity canal to Lamar gage) (Figure 3.5b). Including a trend covariate at subregion 3 yielded a better DIC score (Table 3.4). Reservoir storage and depth to groundwater are not shown in Table 3.4 because each covariate did not improve the DIC score at a

given subregion. Reservoir storage and depth to groundwater are not predictive of monthly NDVI despite the hydrologic connections.

Table 3.4: A summary of the predictive covariates used at each subregion. The autocorrelation coefficient and the variance are not shown in the equations for clarity.

Subregion	Predictive Covariates	DIC
1	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{P}_t + \beta_3 \mathrm{P}_{t-1} + \beta_4 \mathrm{P}_{t-2}$	-1067
2	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{P}_t$	-1012
3	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{t}_t + \beta_2 \mathrm{T}_t + \beta_3 \mathrm{P}_t + \beta_4 \mathrm{P}_{t-1} + \beta_5 \mathrm{P}_{t-2}$	-1063
4	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t$	-760
5	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t$	-793
6	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{P}_t$	-829
7	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t$	-759
8	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{P}_t + \beta_3 \mathrm{P}_{t-1} + \beta_4 \mathrm{P}_{t-2}$	-1029
1	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{Q}_{G/L_t} + \beta_3 \mathrm{Q}_{G/L_{t-1}} + \beta_4 \mathrm{Q}_{G/L_{t-2}}$	-1068
2	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{Q}_{G/L_t} + \beta_3 \mathrm{Q}_{G/L_{t-1}}$	-1025
3	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{t}_t + \beta_2 \mathrm{T}_t + \beta_3 \mathrm{Q}_{G/L_t} + \beta_4 \mathrm{Q}_{G/L_{t-1}}$	-1063
4	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{Q}_{G/L_{t-1}} + \beta_3 \mathrm{Q}_{G/L_{t-2}}$	-761
5	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t$	-795
6	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{Q}_{G/L_{t-1}} + \beta_3 \mathrm{Q}_{G/L_{t-2}}$	-839
7	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t$	-757
8	$\mathrm{y}_t = \beta_0 + \beta_1 \mathrm{T}_t + \beta_2 \mathrm{Q}_{G/L_t}$	-1014

3.3.3 Hypothesis 3: Confluences With Perennial Streams Increase Vegetation Density

The marginal posterior density plots for key spatial model parameters can be seen in Figure 3.6; this graph validates the third prediction. For the average subregion investigated (e.g., average number of confluences with the subregion-specific variance structures ignored for simplicity), areal segments that are intersected with perennial streams have larger temporally-integrated NDVI. The subregions that have a confluence with a perennial stream have increased integrated NDVI compared to subregions with an increased number of confluences and subregions abutted by canal augmentation, floodplain activity, and permanent irrigated land dry-up (Figure 3.6). Subregions abutted by a city, pastoral practices, and permanent irrigated land dry-up have decreased

temporally-integrated NDVI. The spatial autocorrelaed error is smaller than the independent error at all subregions as seen in Figures 3.7 and 3.8. The unstructured error variance is large because the variable encompasses the error between temporally-integrated NDVI for different plant functional groups and densities found in the study area.

The derived quantities are $\bar{\boldsymbol{x}}_a'\beta$, $\bar{\boldsymbol{x}}_b'\beta$, $\bar{\boldsymbol{x}}_c'\beta$, and $\bar{\boldsymbol{x}}_d'\beta$ (where $\bar{\boldsymbol{x}}_a' = (1, \bar{x}_2, 1, 0, 0, 0)'$, $\bar{\boldsymbol{x}}_b' = (1, \bar{x}_2, 0, 1, 0, 0)'$, $\bar{\boldsymbol{x}}_c' = (1, \bar{x}_2, 0, 0, 1, 0)'$) and $\bar{\boldsymbol{x}}_d' = (1, \bar{x}_2, 0, 0, 0, 1)'$). These derived quantities are the average integrated NDVI for a riparian zone that is abutted by either augmentation or cities and pastoral practices or intersected by a perennial river or abutted by permanent irrigated land dry-up. The mean and 95% central intervals (2.5% and 97.5%) for these derived quantities can be seen in Figure 3.6. The contrast vector $\bar{\boldsymbol{x}}_c'\beta$ indicates that subregions that have a perennial confluence have noticeable larger temporally-integrated NDVI while $\bar{\boldsymbol{x}}_b'\beta$ (floodplain activity) and $\bar{\boldsymbol{x}}_d'\beta$ (land dry-up) have a decreasing impact on temporally-integrated NDVI.

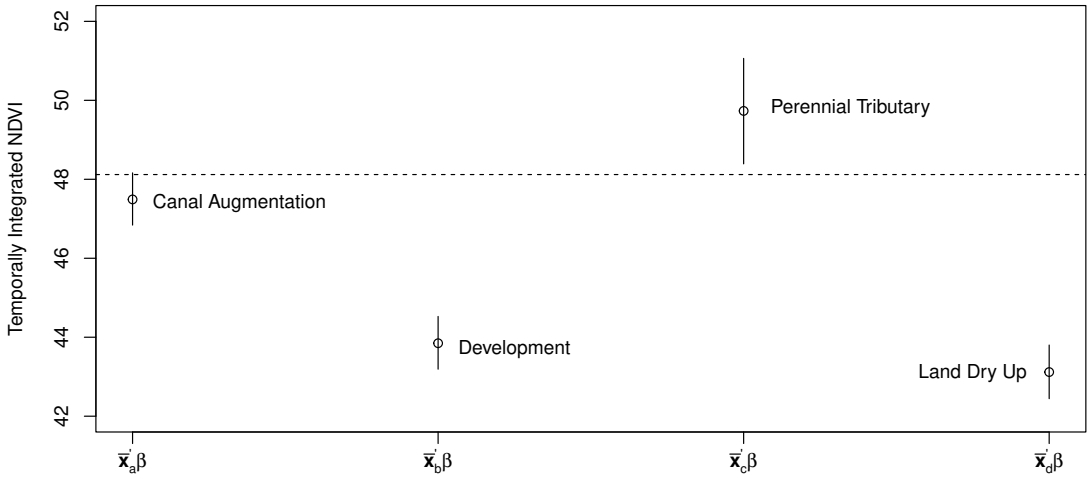

Figure 3.6: Marginal posterior means (open-circle points) and 95% credible intervals (vertical lines) for the computed contrasts. The horizontal-dashed line represents the posterior mean of β_0.

3.4 Discussion

This work using statistical analysis on publicly available water-resources data can be used to help resource managers understand lagged catchment processes in perturbed systems. Before this

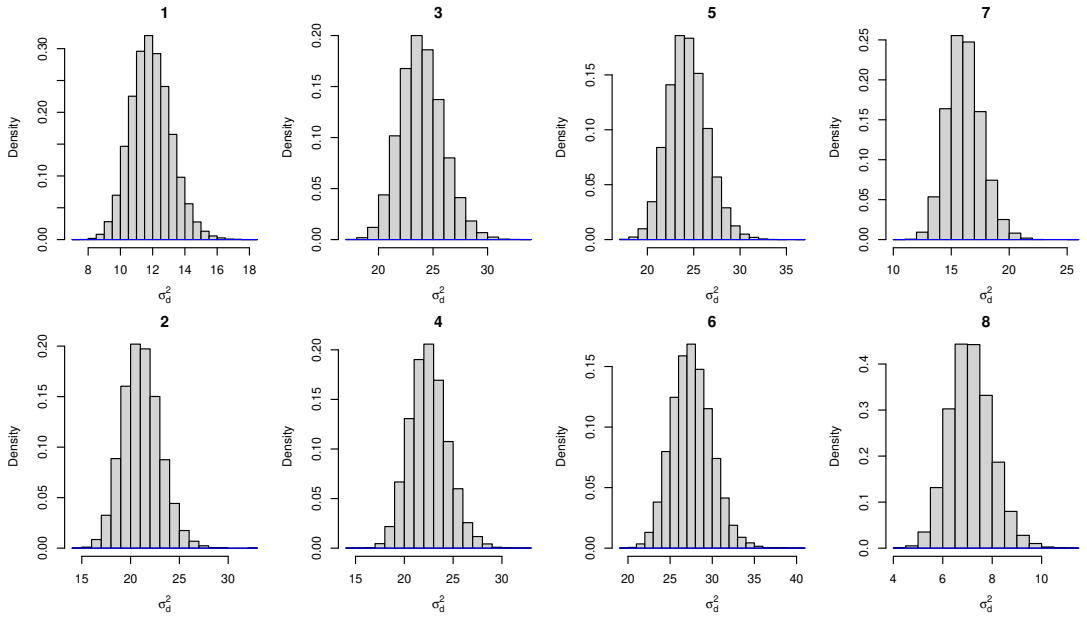

Figure 3.7: Marginal posterior densities of the unstructured error variance (σ_d^2) per subregion.

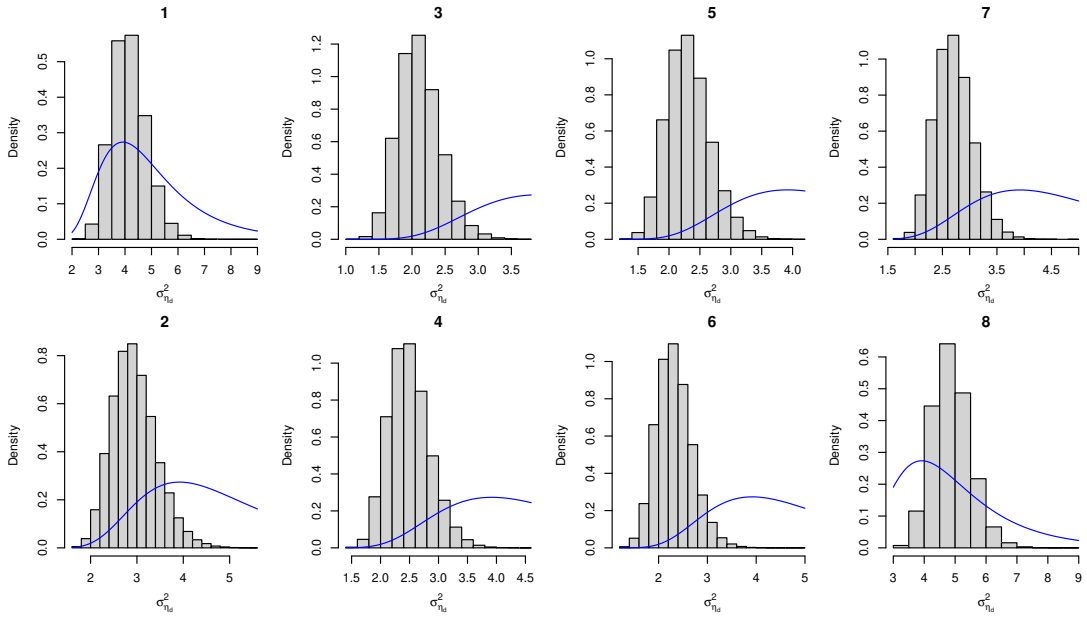

Figure 3.8: Marginal posterior densities of the spatially-correlated error variance ($\sigma_{\eta_d}^2$) per subregion.

work is synthetically placed in the broader context, there are a few key drawbacks to this study to consider. The spatial resolution of the Landsat NDVI product is coarse. The coarseness of the imagery makes inference on a specific vegetation type difficult in a heterogeneous riparian corridor. I attempt to overcome that in the spatial approach by accounting for correlated (i.e., neighbor) and

uncorrelated (i.e., random) error variance. In the temporal and spatial analysis, the data model is Gaussian which has support from $-\infty$ to ∞ and NDVI is strictly positive. This drawback is quickly overcome using conjugate relationships between the data, process, and parameters.

3.4.1 Hypothesis 1: Identifiable Temporal Trend At Catchment Scale

A growth trend is discovered in riparian NDVI at spatial resolutions less than catchment scale. Subsequently, in terms of the strategic areas that are examined, the subregion from the Amity canal to the Lamar gage is the only subregion that shows a significant temporal trend in monthly NDVI as indicated by the posterior histogram that does not significantly overlap zero (95% credible interval); as summarized in Figure 3.5b. Seven out of eight subregions did not exhibit a significant slope of growth or decay in monthly NDVI. In contrast, Nguyen et al. (2015) found a clear-decreasing trend using linear least-squares analysis on mean annual values of riparian vegetation indices for a 39.51 km^2 plot on the upper San Pedro River in Arizona. Figure 3.5 also summarizes the marginal posterior distributions for the linear regression coefficients that represent the intercept, temperature and the autocorrelation coefficient for the AR(1) term (i.e., the baseline model). All subregions in the study area showed a strong first order autoregressive error (Figure 3.5d). The relationship between the responses and covariates at time t and $t-1$ create the autoregressive structure and the mean AR(1) value for all subregions combined is approximately 0.5. Temporal dependence in a water resource time series like the data analyzed in this study can be caused by atmospheric and hydrologic forcings, as well as agriculture and pastoral practices [83, 106].

3.4.2 Hypothesis 2: Precipitation Will Have A Shorter Lag Effect

The correlation of a two-month lag, or less, between precipitation and riparian NDVI is likely due to the area receiving most of its rainfall during the summer months and by spring snow replenishing soil moisture and recharging the aquifer in time for riparian green-up [25]. Nguyen et al. (2015) found direct correlations between generalized NDVI and precipitation. Vanderhoof et al. (2019) tested the significance of lagged precipitation to explain the temporal changes in riparian wetness and found that annual precipitation with drought indices are good predictors of

the observed decline of riparian wetness. Parsons and Thoms (2013) found discernible greenness changes in riparian vegetation during periods of increased rainfall and flooding [107]. Huntington et al. (2016) measured moderate, but direct correlations between summer NDVI and annual precipitation and groundwater depth.

In a time series approach, the variable Q_{G/L_t} is a proxy for when return flow is causing a stream segment to gain or lose. The delayed correlation between monthly NDVI and Q_{G/L_t}, as shown in Table 3.4, highlights the spatial and temporal variation in ecosystem function along the Arkansas River. Previous studies like Nagler et al. (2016) have used river discharge loss as a way to validate riparian water use for approximately 380 km^2 of red gum riparian forests in the Yanga National Park, Australia [108]. Velpuri et al. (2020) used estimates of satellite derived evapotranspiration and river discharge to quantify how irrigation curtailment impacted the water balance in the upper Klamath Lake basin [109]. This work compliments previous efforts to relate stream mass-balance to floodplain activity by examining the lags.

The lack of correlation between riparian NDVI and groundwater depth is likely due to the sparseness of groundwater data. Most groundwater observations were recorded twice per year in the spring and fall to capture static water levels at pre- and post-irrigation time stamps. The lack of correlation between groundwater and NDVI, in addition to reservoir activity and NDVI, does not agree with previous groundwater dependent ecosystem studies [7] and research on the benefits of environmental flows [110].

3.4.3 Hypothesis 3: Confluences With Perennial Streams Increase Vegetation Density

Quantifying the connection of irrigating communities and riparian water use remains an elusive component of the hydrosocial framework. Spatial autoregressive statistics, like the methodology used here, is required when the data has spatial dependence and has been used to detect land use change from driving factors [111]. Based on the spatial analysis, a subregion that has a confluence with a perennial stream (e.g., Mud, Big Sandy and Wild Horse Creek) has increased integrated

NDVI when compared to sites that are impacted by augmentation, floodplain activity or permanent irrigated land dry-up (Figure 3.6). Albano et al. (2020) analyzed the residuals of a vegetation-climate regression analysis in relation to non-climatic forcings in Nevada and found, outside of the model, negative temporal trends in riparian vegetation were abutted by agriculture. Albano et al. (2020) found growth in the temporal patterns of riparian vegetation that were connected to tributaries and perennial ponds and wetlands. Vanderhoof et al. (2019), outside of the random forest regression, related the trend in the residuals of a climate-vegetation model fit to changing irrigation methodologies over a given area. The spatial knowledge acquired from this work could be used to inform strategic field efforts aimed at preserving vulnerable ecosystems.

3.5 Conclusion

With the use of temporal and spatial statistics, this work provided an alternate way of validating previous observations and modeling that highlight floodplain connections to riparian ecosystems in groundwater-dependent regions. This work discovered that: (a) temporal trend is spatial-scale dependent, (b) precipitation has a shorter lagged impact on riparian NDVI when compared to other hydrologic factors despite studying a groundwater dependent ecosystem, and (c) that perennial tributary confluences explain increased vegetation density. Future efforts in this direction will consider combining the independent analyses used in this investigation into a single spatio-temporal approach similar to [112]. There is opportunity to link well-known models for handling spatial data with those for handling latent temporal patterns [113].

Chapter 4

Point-Scale Relationship Between Riparian Evapotranspiration and Groundwater

4.1 Introduction

Natural-ecosystem evapotranspiration (ET) is a difficult-to-quantify flux of the hydrologic cycle and the relationship to groundwater dependent ecosystems (GDE) has important implications on the co-management of groundwater and surface water supply in Colorado. Germane to this topic is the point-scale deterministic fit between ET and groundwater depth; and how this relationship depends on plant functional group. Literature has shown that change point (i.e., threshold or piecewise) models are adequate predictors of the relationship between ET and groundwater depth but there is observational support that the relationship between groundwater depth and plant health approach saturating conditions (i.e., asymptotic nature) in shallow groundwater. A stochastic framework to derive the non-linear or linear deterministic fit between ET and groundwater head in GDEs is overdue. For the purposes of creating scale-enlightened models, it is imperative to test the predictive abilities of these function forms.

Understanding the mechanistic process between evapotranspiration and groundwater head, and the scale at which the process occurs, is key to building models that are predictive at scales relevant to regional water challenges. There are numerous groundwater modeling schemes that heavily rely on the assumption of linear and piecewise-linear between evapotranspiration and groundwater head. For instance: 1) MODFLOW is the U.S. Geological Survey's internationally recognized modular hydrologic model and MODFLOW has an Evapotranspiration Package [114] and Evapotranspiration Segments Package [115] that simulate ET rates as a function of hydraulic head in the saturated zone using linear and piecewise linear functions, respectively (Figure 4.1); 2) Lumped, or bucket, models describe the relationship between evapotranspiration and water table depth as

piecewise linear [19]; and 3) variable saturation flow models (1-D Richard's equation) have been used to show the usefulness of a piecewise linear process between ET and groundwater [116].

The assumption of linearity between ET and groundwater is valid at coarse spatial and temporal scales because it is well known that as groundwater depth decreases (i.e., shallowing), and comes in contact with roots, then evapotranspiration increases. At smaller scales, the assumption of linearity might be inadequate due to vegetation heterogeneity in riparian corridors. Water use by phreatophytes, shrubs, and grasses may all have different mechanistic relationships (e.g., ET rates) with groundwater. Different deterministic equations may serve as predictive representations of the ET-groundwater process that depend on plant functional group.

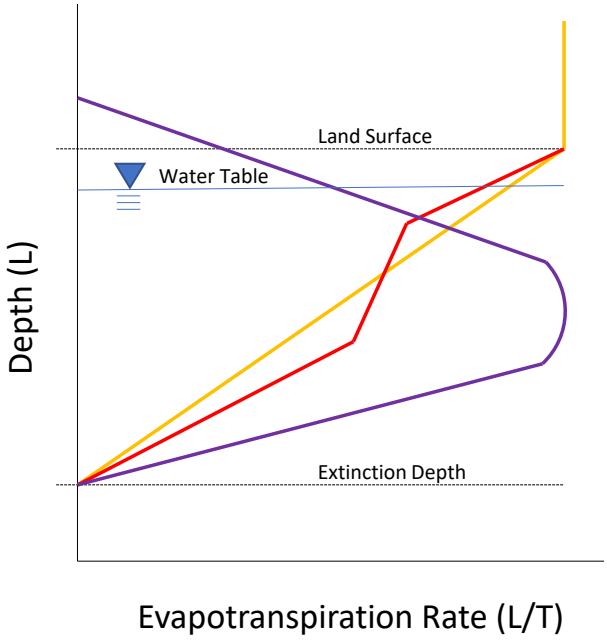

Figure 4.1: Conceptualization of the functional relationship between groundwater depth and evapotranspiration in MODFLOW. The yellow line is linear, the red line is piecewise-linear, and the purple line is hyperbolic and each line represents a different MODFLOW module for simulation the ET-groundwater relationship.

Previous literature suggests there is observational evidence for a linear and non-linear relationship between evapotranspiration and groundwater. Specifically, a linear fit has been found using only groundwater as a covariate [13, 114, 117, 118], exponential relationships between ET and

groundwater depth have been observed [13, 118, 119], piecewise exponential-linear equations have been derived [116, 120–122], piecewise linear [115], hyberbolic [123], and asymptotic (i.e., saturating) conditions have been observed between normalized difference vegetation index (NDVI) and groundwater depth observations [7].

Because natural ecosystem evapotranspiration is more difficult to accurately estimate when compared to crops, lacking uniform vegetation type and fetch, research on natural ecosystems have used ratios of evapotranspiration. Ratios of evapotranspiration are synonymous with crop coefficients and consist of measuring ET in a specific location and dividing that value by a climatic potential. The FAO-56 documentation quantifies ET in non-pristine vegetation using the ratio of actual ET to potential ET; these values of ET are quantified using the Penman-Montieth equation which has been validated in Colorado [63]. Ratios of these variables can convert an estimate of water use (mm/day) into a value of water use with a limit [124]. Previous researchers have computed ratios of water use as a response variable [13, 105, 116, 125]. Normalizing an evapotranspiration rate by its climatic potential is common practice in large-scale hydrology [126, 127]. While there is observational support for linear and non-linear relations between evapotranspiration and groundwater depth, there has not yet been a study to examine the predictive ability of linear and non-linear functions in a Bayesian framework at varying temporal resolutions. Understanding which function is most predictive of ET using groundwater depth could improve spatially-distributed parameter models with built in ET-groundwater understanding. Additionally, Bayesian inference on process parameters could provide ecohydrologic understanding.

The objective of this work is to derive the functional fit between evapotranspiration and groundwater depth for three plant functional groups (PFG) found in arid GDEs. The three PFGs are shrub, phreatophyte, and grass. The broader importance of this work is aimed at understanding linkages between groundwater and ecosystem processes to sustainably manage interconnected groundwater and surface water systems in an irrigated agricultural region.

The hypothesis I have about GDEs is formulated as follows: asymptotic functions are hypothesized to serve as more predictive models of the relationship between GDE health (i.e., evapotran-

spiration) and groundwater availability (i.e., groundwater hydraulic head) given that no quantities in hydrology can increase or decrease without limit. Testing this hypothesis will involve using a Bayesian regression framework. The framework will facilitate the testing and predictive analysis for various function forms used in ecohydrology. The hypothesis will be proved if an asymptotic function provides the best fit between groundwater head and ET. Results will be presented on the predictive abilities of each function using within sample, and out-of-sample, model validation at varying temporal resolutions for three PFGs.

4.2 Methods

4.2.1 Study Area

The field study was conducted near the CSU Arkansas Valley Research Center in Rocky Ford, CO at a riparian-wetland combination GDE that borders a 5-km stretch of the Arkansas River (Figure 4.2). The spatial domain includes a low-gradient GDE that is surrounded by the Rocky Ford State Bird Farm and fallow land that is predominantly owned by upstream municipalities along the front range of the Rocky Mountains. This study site is contained in a dryland setting where precipitation is typical of arid climates [128], yet exhibits humid land aspects because of the shallow groundwater depths [129]. Data from the micrometeorology station, *RFD01* (Lat: $38.0385°$, Long: $-103.6950°$), operated by the Colorado Agriculture Meteorological Network (CoAgMet), indicates that water year (Oct.-Sept.) precipitation totals are 158.74-mm (WY-2018), 261.11-mm (WY-2019), and 242-mm (WY-2020). The vegetation in the study area is in a transitional phase to an alternative stable state due to succession from competitive vegetation (e.g., Tamarix ramnossima) [59, 130]. The combination of deciduous trees and desert shrub, shallow groundwater and strong atmospheric drying power make this study area a fascinating ecological paradox.

4.2.2 Field Data

Point-scale, 15-minute, observations of groundwater table elevation at 12 monitoring wells located within, and adjacent to, the riparian corridor along the Arkansas River were collected. The

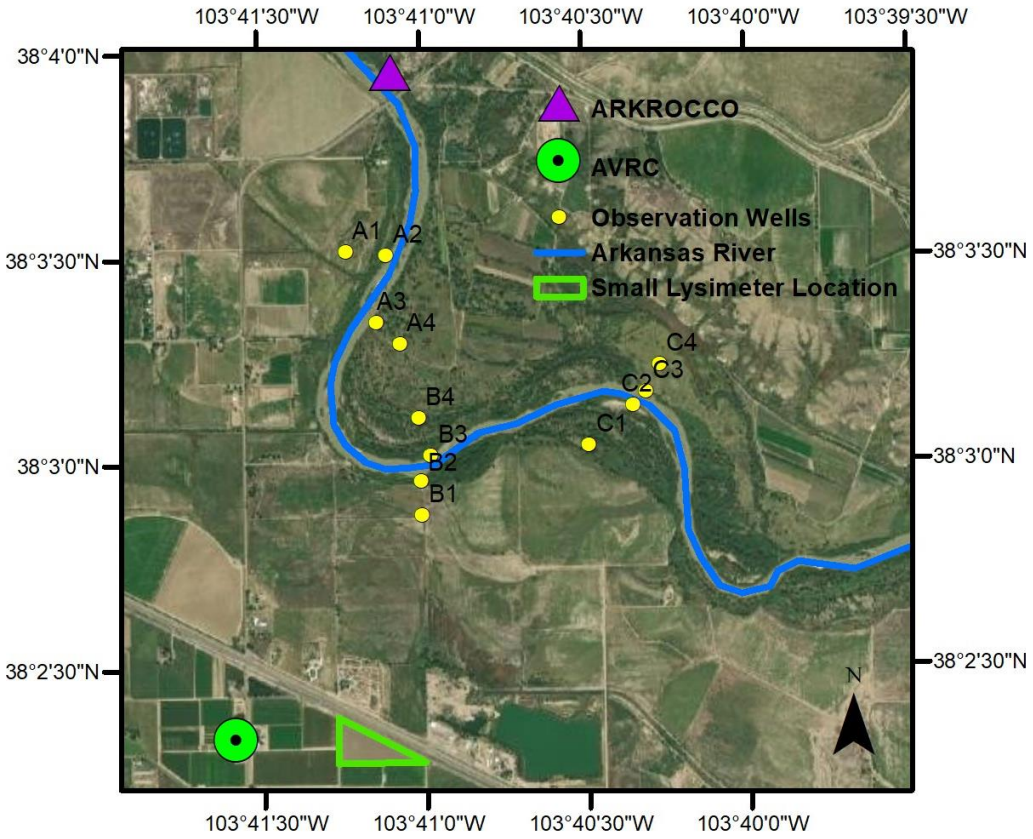

Figure 4.2: General location map of the study area in Rocky Ford, CO.

groundwater wells (HOBO Water Level Data Logger) were installed in 2014 and have been previously used to monitor stream-aquifer processes with the purpose of better understanding temporal variability in Nitrogren transport [131]. Groundwater measurements at all monitoring wells began May 15, 2018 and ended on October 17, 2020. Water table elevation was calculated at the well by subtracting the barometric pressure from the total pressure measured by the pressure transducer in each monitoring well. Each well was geo-located with survey-grade TOPCON Real-Time Kinematic GPS as the base station that is paired with TOPCON GR-5 rovers to collect latitude, longitude, and elevation with a 10 mm horizontal error and 15 mm vertical error.

Point-scale, 15-minute averaged, micrometeorological data within the riparian corridor using a HOBO Onset weather station located near groundwater well A2 (38.0582°, -103.6852°) were collected. The choice of this location was determined by ease of access, landowner clearance, and to capture the GDE micrometeorology. Daily data from the RFD01 CoAgMet station, located at

the CSU Arkansas Valley Research Center in Rocky Ford, CO, were downloaded from the CoAgMet website. Daily values of ET, within a plot of salt cedar and at RFD01, are computed using the Penman-Monteith equation (Equation 4.1) following the methodology outlined in the ASCE-EWRI Task Committee Report [60]. Standardized evapotranspiration (ET_{sz}) from the weather station in the riparian corridor is divided by grass-reference evapotranspiration (ET_o) measured by the RFD01 station to create a normalized value of evapotranspiration (ET_{sz}/ET_o). The names, units, and how each variable is computed in Equation 4.1 is summarized in Table 4.1. The only difference in computing ET_{sz} compared to ET_o is the numerator (C_n) and denominator (C_d) constants which are based on time step, aerodynamic roughness, and bulk surface resistance. Per Table 4.1 in the ASCE-EWRI Task Committee Report, ET_o is computed using the *short* reference values and ET_{sz} is computed using the *tall* reference values. Additionally, the wind profile is measured at a height of three meters at the weather station located in the GDE, but the profile is corrected to a standard two-meter measurement following Equation 33 in ASCE-EWRI Task Committee Report.

$$ET_{sz} \text{ and } ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{C_n}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + C_d u_2)} \tag{4.1}$$

4.2.3 Remote Sensing Data

A collection of 276 enhanced multispectral visible and near-infrared CubeSat images from a commercial platform called Planet (www.planet.com), as part of the Education and Research Program, are used in this study. Planet's PlanetScope sensors provide a daily surface-reflectance product that is nadir-pointing ($< 5°$) with a ground-sampling distance of 3.125 m [18]. The surface-reflectance image from PlanetScope Ortho Tiles has radiometric and sensor corrections applied to the data. Each image is orthorectified and projected to Universal Transverse Mercator (UTM) Zone 13N. To retrieve the imagery from Planet, the following steps are completed: 1) project the study-area shapefile to Planet's preferred coordinate reference (UTM Zone 13N, WGS 1984), 2) convert the study-area shapefile to geojson format using geojson.io, and 3) then send a download request through an API platform called Postman. The normalized difference vegetation

Table 4.1: Input data from weather station to compute tall-reference evapotranspiration ($ET_{rs} = ET_{sz}$). All variables were computed using the ASCE-EWRI Task Committee Report (2005).

Variable Name	Units	Reference
Numerator constant (C_n)	$Kmms^3Mg^{-1}d^{-1}$	Table 1
Denominator constant (C_d)	sm^{-1}	Table 1
Mean air temperature (T)	$^\circ C$	Eqn. 2
Atmos. pressure (P)	kPa	Eqn. 3
Psychrometric constant (γ)	$kPa^\circ C^{-1}$	Eqn. 4
Slope of saturation vapor press.-temp. curve (Δ)	$kPa^\circ C^{-1}$	Eqn. 5
Net radiation at the crop surface (R_n)	$MJm^{-2}d^{-1}$	Eqn. 15
Saturation vapor pressure (e_s)	kPa	Eqn. 6
Actual vapor pressure (e_a)	kPa	Eqn. 11
Incoming solar radiation (R_s)	$MJm^{-2}d^{-1}$	Measured
Net short-wave radiation (R_{ns})	$MJm^{-2}d^{-1}$	Eqn. 16
Net long-wave radiation (R_{nl})	$MJm^{-2}d^{-1}$	Eqn. 17
Clear-sky solar radiation (R_{so})	$MJm^{-2}d^{-1}$	Eqn. 19
Extraterrestrial radiation (R_a)	$MJm^{-2}d^{-1}$	Eqn. 21
Soil heat-flux density (G)	$MJm^{-2}d^{-1}$	Eqn. 30
Wind speed (u_2)	ms^{-1}	Eqn. 33

index (NDVI) is computed for each image using the well-known ratio of near-infrared and red reflectances shown in Equation 4.2. The NDVI product has been validated for usage on groundwater dependent ecosystems [91].

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{4.2}$$

4.2.4 Sampling Locations

Figure 4.3 displays 12 groundwater wells, lysimeter location, and each sampling location to sample NDVI on the day of satellite overpass. Each sampling area is 529 m^2 to capture > 50 image pixels to describe a single plant functional group. Prior field experience, latest National Agriculture Imagery Program imagery, and peak-season NDVI are all used to ensure that each quadrat included one PFG where possible.

Planet Imagery satellites are known as CubeSats, which are new to the ecohydrologic research community [18]. These satellites are ultrahigh resolution and offer the ability to gain new insight on ecohydrological systems and processes but need to be ground-truthed before being used

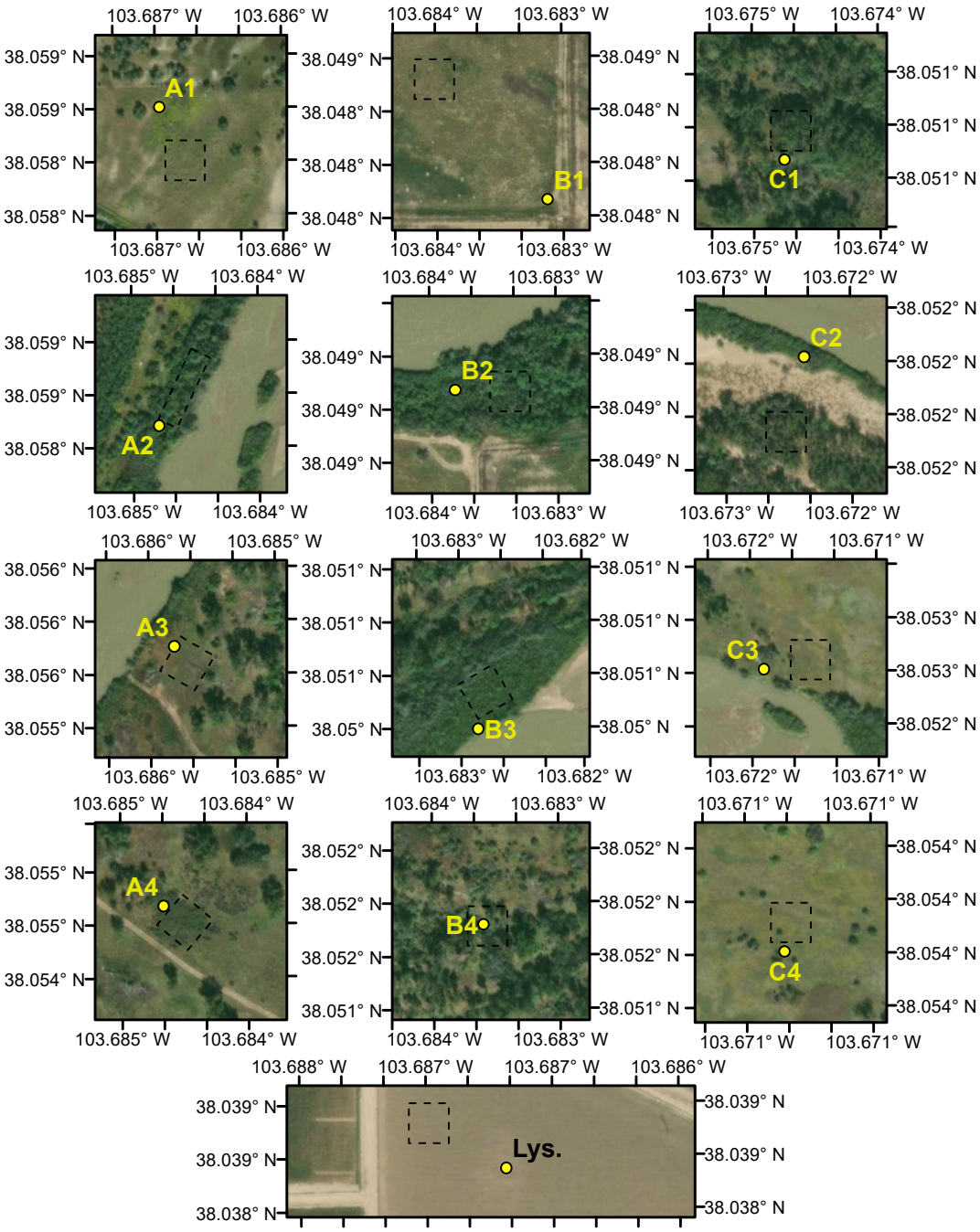

Figure 4.3: Location map for all 12 groundwater monitoring wells, small-weighing lysimeter, and the corresponding vegetation sampling location.

to make inference. Therefore, Planet NDVI is compared to mass-based evapotranspiration to increase confidence in using the product as a proxy for natural ecosystem evapotranspiration. The CSU Arkansas Valley Research Center has two precision weighing lysimeters that were in opera-

tion during the study period [62]. Data from the small-weighing lysimeter is used in this analysis because the crop was consistently grass plantings during the study period. The daily lysimeter data includes estimations of evapotranspiration (ET_c) and a list of observed events (e.g., rain, irrigation, snow, lysimeter drain). Data was filtered from the small-weighing lysimeter by removing all data from the 2018 growing season, after the second cutting (i.e., harvest) each growing season, removing one week of data after the first cutting of each growing season, and removing days when a non-standard event occurred. There is agreement between Planet NDVI and lysimeter ET_c as shown in Figure 4.4. Reduced correlation could be due to the fact that NDVI only accounts for the transpiration portion of ET. This strict mass-based measurement of ET_c validates the use of Planet NDVI as a proxy for plant water use in the study area.

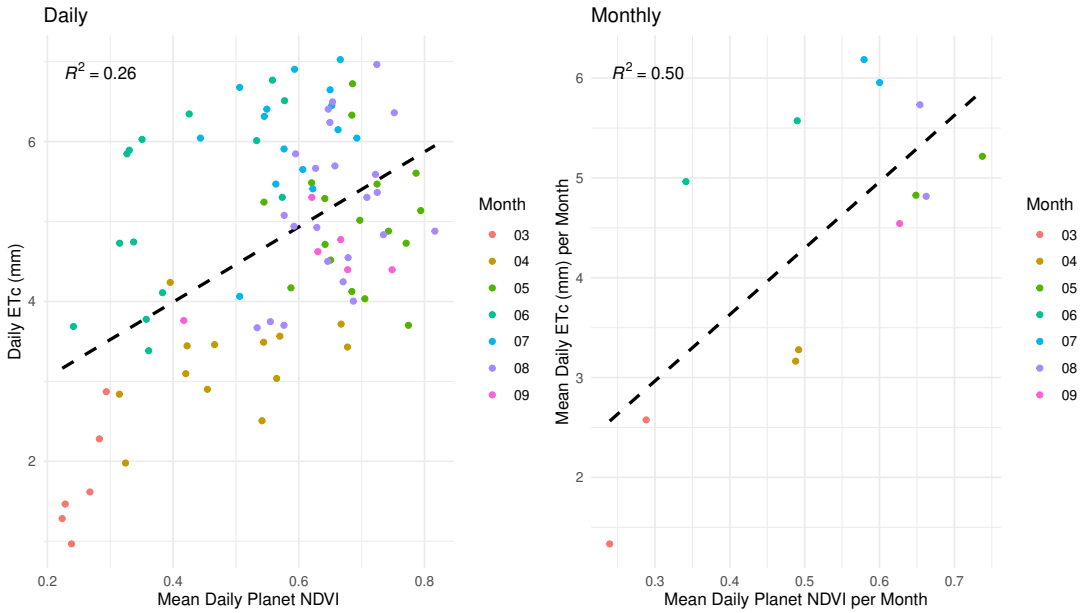

Figure 4.4: Linear regression between daily (left) and monthly (right) Planet NDVI and ETc from the small-weighing lysimeter.

4.2.5 Prediction: Asymptotic Functions Are More Predictive Than Linear

The data is prepared on a daily and monthly time scale for the regression analysis. Daily standardized evapotranspiration (ET_{sz}) and grass-reference evapotranspiration (ET_o) is converted

to a ratio by dividing both daily values on each day of the temporal domain of the study. Daily values are summed for each month, for each variable (ET_{sz} and ET_o), and divided to create a monthly ratio of water use. The 15-minute values of hydraulic head are converted to daily values (meters above sea level) by computing the arithmetic average of all values on a given day. The same process to generate daily values of hydraulic head (H_{masl}) is used to compute monthly values. Figure 4.5 displays the non-linear relationship between ET and groundwater on a daily and monthly time scale for three plant funcational groups. Figure 4.6 displays a similar non-linear relationship, from a biophysical perspective, between NDVI and groundwater on a daily and monthly time scale. The data shown in Figure 4.5 is used as direct input into the regression analysis designed to derive the function of best fit.

The regression analysis is explained by discussing key details of the model setup for the three models (*control, asymptote,* and *change point*, as shown in Table 4.2) and how the models are scored. Appendix B shows the model posterior distribution for all three models (Section B.1) and the full conditional statements for all model variables (Section B.2). All models are non-hierarchical which assumes that measurements of evapotranspiration are made without error. A custom Markov chain Monte Carlo (MCMC) sampler is created for the models shown in Table 4.2 following the six general steps outlined by Hobbs and Hooten (2015) [89]. Each model in Table 4.2 has an intercept and groundwater head is the only covariate. The covariate is standardized before being inserted into each model to speed convergence [89]. All models are tested with simulated data (i.e., data generated with known model-parameter values) to confirm that each algorithm is capable of reproducing known model variables [132]. All starting values for each model are obtained from point estimates by computing summary statistics on the data. The number of iterations in each MCMC simulation is $k = 100,000$ with a burn-in period of $20,000$. Model convergence is confirmed through visual inspection of trace plots. Mean square prediction error (MSPE) is used as a score function for each model (Equation 4.3 and 4.4) where $\hat{y}_i$ is the model prediction of the ith observation in the dataset. Low values of MSPE indicate increased predictive ability [89].

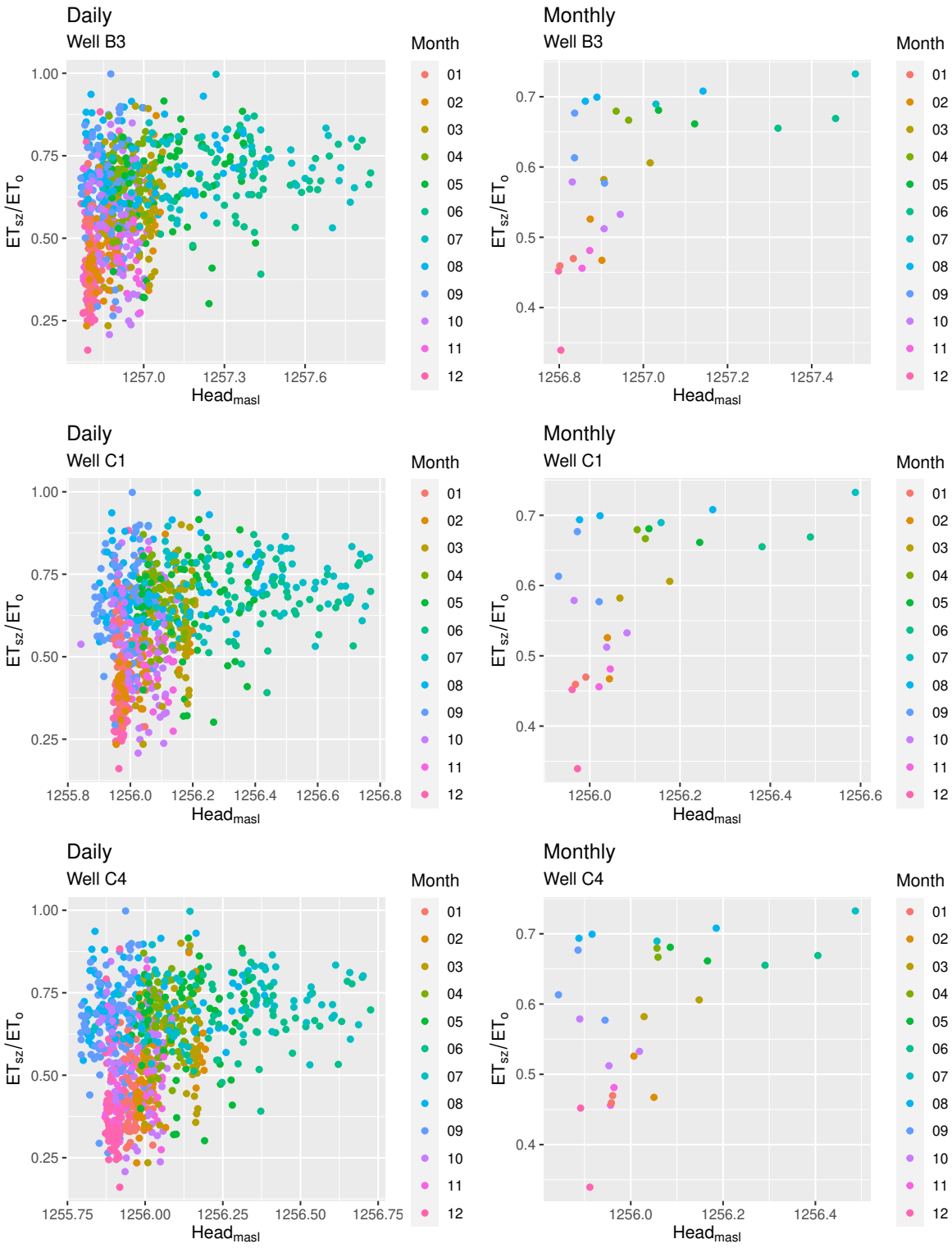

Figure 4.5: Observations of non-linear behavior in the groundwater-ET relationship at observation well B3 (shrub,saltcedar), C1 (phreatophyte, cottonwood), and C4 (grass, prairie grass) on a daily and monthly time scale.

Figure 4.6: Observations of non-linear behavior in the groundwater-NDVI relationship at observation well B3, C1, and C4 on a daily (top) and monthly (bottom) time scale.

$$\hat{\mathrm{y}}_i = \frac{\sum_{k=1}^{K} \mathbf{g}(x)^{(k)}}{\mathrm{K}} \tag{4.3}$$

$$\mathrm{MSPE} = \frac{\sum_{i=1}^{N} (\mathrm{y}_i - \hat{\mathrm{y}}_i)^2}{\mathrm{N}} \tag{4.4}$$

4.2.6 Prior Information

Objective prior information on model variables was used with a few exceptions. Each model begins with a normal distribution to portray the response variable where y_i represents the i^{th} observation of ET_{sz}/ET_o from $i = 1, ..., N = 820$ for the daily analysis and $i = 1, ..., N = 27$ for the monthly analysis.

The control model is a classic linear regression model in a Bayesian setting with only one covariate ($x_{i,1}$) which is standardized groundwater head at the i^{th} observation. The control model is a fully Gibbs MCMC algorithm [79]. For the control model: β_0 is a normal scalar regression coefficient for the intercept with a mean of μ_{β_0} and variance of $\sigma^2_{\beta_0}$, β_1 is a normal scalar regression coefficient for the groundwater-head covariate with a mean of μ_{β_1} and variance of $\sigma^2_{\beta_1}$, and the process variance, σ^2_y, is represented with an inverse gamma distribution with a shape of q_y and scale of r_y.

The asymptote model is a hybrid Gibbs and Metropolis-Hastings algorithm and the function form is based on the Michaelis-Menten equation which was originally developed to model enzyme-catalysed chemical reactions [89, 133]. This formulation is used because the model variables are easy to interpret in a ecohydrological context as opposed to a negative exponential equation. For the asymptote model: the v_{max} variable is a scalar defined as the maximum value of the independent variable with a prior mean (μ_v) of 0.75 and a vague variance (σ^2_v) which is informed by a study of riparian ET along the Platte River in Nebraska [13], the a variable is a scalar defined as the rate of change in ET_{sz}/ET_o at deeper values of groundwater head with a prior mean (μ_a) and variance (σ^2_a), the c variable is a scalar that acts as a horizontal shift (i.e., transformation) with a mean (μ_c) and variance (σ^2_c), and the process variance, σ^2_y, is represented by an inverse gamma distribution with a shape of q_y and scale of r_y. All tuning variables ($\sigma^2_{v,tune}$, $\sigma^2_{a,tune}$, $\sigma^2_{c,tune}$) are iteratively determined to achieve an acceptance percentage of 10-60$\%$ with each Metropolis-Hastings sample.

The change point model is non-linear but the formulation is a combination of two linear functions with a threshold value to indicate when vegetation transitions from dormant to actively growing. The change point model is a fully Gibbs MCMC algorithm. The $x_{i,1}$ is standardized groundwater head at the i^{th} observation and $x_{i,2}$ is an indicator variable that is 0 if $x_{i,1} \leq cp$ and 1 if $x_{i,1} > cp$. For the change point model: the cp variable is a scalar that represents a change in ecological state (i.e., end of dormancy) that happens in April-May [25, 59] and is represented with a prior mean of μ_{cp} and a variance σ^2_{cp}, there is no literature supporting values of the intercept (β_0) so the variable has a sufficiently vague prior mean (μ_{β_0}) and variance ($\sigma^2_{\beta_0}$), β_1 is a normal scalar

regression coefficient for the groundwater-head covariate that represents the rate between ET and groundwater head prior to the change in ecological state with mean of μ_{β_1} and a variance of $\sigma^2_{\beta_1}$, β_2 is a normal scalar regression coefficient for the groundwater-head covariate that represents the rate between ET and groundwater head after the change in ecological state with mean of μ_{β_2} and variance of $\sigma^2_{\beta_2}$, the prior means of the rates (β_1 and β_2) are informed with positive slopes following several studies [50, 114, 115], and the process variance, σ^2_y, is represented by an inverse gamma distribution with a shape of q_y and scale of r_y.

Table 4.2: A summary of the models used in this study. The symbol $\sim$ implies two sides of an equation are proportional to each other. The probability distribution abbreviations are as follows: normal-Gaussian (N), and inverse gamma (IG).

Control	Asymptote	Change Point
$\mathrm{y_i} \sim \mathrm{N}(g(x)_i, \sigma^2_\mathrm{y})$	$\mathrm{y_i} \sim \mathrm{N}(g(x)_i, \sigma^2_\mathrm{y})$	$\mathrm{y_i} \sim \mathrm{N}(g(x)_i, \sigma^2_\mathrm{y})$
$g(x)_i = \beta_0 + \beta_1 \mathrm{x}_{i,1}$	$g(x)_i = \frac{\mathrm{v}_{max} \times (\mathbf{x}_i + \mathrm{c})}{\frac{\mathrm{v}_{max}}{a} + (\mathbf{x}_i + \mathrm{c})}$	$g(x)_i = \beta_0 + \beta_1 \mathrm{x}_{i,1} + \beta_2 (\mathrm{x}_{i,1} - \mathrm{cp}) \mathrm{x}_{i,2}$
$\beta_0 \sim \mathrm{N}(\mu_{\beta_0}, \sigma^2_{\beta_0})$	$\mathrm{a} \sim \mathrm{N}(\mu_\mathrm{a}, \sigma^2_\mathrm{a})$	$\beta_0 \sim \mathrm{N}(\mu_{\beta_0}, \sigma^2_{\beta_0})$
$\beta_1 \sim \mathrm{N}(\mu_{\beta_1}, \sigma^2_{\beta_1})$	$\mathrm{v}_{max} \sim \mathrm{N}(\mu_\mathrm{v}, \sigma^2_\mathrm{v})$	$\beta_1 \sim \mathrm{N}(\mu_{\beta_1}, \sigma^2_{\beta_1})$
$\sigma^2_\mathrm{y} \sim \mathrm{IG}(q_y, r_y)$	$\mathrm{c} \sim \mathrm{N}(\mu_\mathrm{c}, \sigma^2_\mathrm{c})$	$\beta_2 \sim \mathrm{N}(\mu_{\beta_2}, \sigma^2_{\beta_2})$
$-$	$\sigma^2_\mathrm{y} \sim \mathrm{IG}(q_y, r_y)$	$\mathrm{cp} \sim \mathrm{N}(\mu_\mathrm{cp}, \sigma^2_\mathrm{cp})$
$-$	$-$	$\sigma^2_\mathrm{y} \sim \mathrm{IG}(q_y, r_y)$

4.3 Results and Discussion

Out of 12 well locations, the results focus on well B3 (shrub), C1 (phreatophyte), and C4 (prairie grass) because each site contains a homogeneous distribution of a singular plant functional group. Additionally, the shrub site contains an invasive species called tamarisk (Tamarix spp., aka saltcedar) and the water use patterns of tamarisk are of particular interest in Colorado and the western United states [70]. The extensive root systems of phreatophytes, making them largely dependent on groundwater to satisfy water use requirements [28], provide a cross-comparison with shallow-rooted grasses.

4.3.1 Prediction: Asymptotic Functions Are More Predictive Than Linear

The predictive ability of the asymptote function is useful in groundwater-dependent ecosystems. Table 4.3 validates the hypothesis by indicating that the asymptote model is more predictive than the control and change point models but it depends on plant functional group and temporal scale. Based on MSPE scores, the asymptote model is more predictive than the control and the change point models on a daily and monthly time scale for the shrub PFG. The change point model is most predictive for the phreatophyte PFG. The control model is equally predictive as the change point model and more predictive than the asymptote model on a daily and monthly time scale for the grass PFG. Figure 4.7 shows how each model fits the within-sample, daily and monthly dataset for three plant functional groups.

Exponential functions are not tested in this analysis because the model parameters are less interpretable than the Michaelis-Menten formulation of the asymptotic function. However, research shows that exponential curves between evapotranspiration and groundwater are highly predictive. For example, Sanderson and Cooper (2008) derived an exponential curve between evapotranspiration as a function of groundwater depth for a grassland ecosystem in an intermountain basin on an annual time scale [125]. Yue et al. (2016) derived an exponential curve between a daily ratio of actual evapotranspiration and potential evapotranspiration against groundwater depth along a riparian corridor of the Platte River in Nebraska [13]. Nichols (1994) found exponential curvature between evapotranspiration and groundwater for separate distributions of phreatophyte shrubs in Nevada [119].

There is limited use of the Michaelis-Menten version of the asymptotic function in groundwater dependent ecosystem modeling. The Michaelis-Menten formulation of the asymptotic function contains variables with ecohydrologic meaning that are easily interpreted. For example, Figure 4.8 provides the derived distributions for each model parameter of the asymptote model for three plant functional groups. The mean of variable a, defined as the rate of change in ET_{sz}/ET_o at deeper values of groundwater depth, is approximately 1.9 (11.5 with an non-standardized covariate) for the shrub PFG. The v_{max} variable represents the maximum value of ET_{sz}/ET_o as groundwater

Table 4.3: A summary of the mean squared prediction error for each model.

Well	Model	Type	Data	MSPE Score
B3	Control	Within	Daily	0.02193228
	Asymptote	Within	Daily	**0.02092503**
	Change Point	Within	Daily	0.02110829
B3	Control	Within	Monthly	0.007064935
	Asymptote	Within	Monthly	**0.005541293**
	Change Point	Within	Monthly	0.005770637
B3	Control	OOS	Daily	**0.02601304**
	Asymptote	OOS	Daily	0.02859786
	Change Point	OOS	Daily	0.02715694
B3	Control	OOS	Monthly	0.006847198
	Asymptote	OOS	Monthly	**0.006036566**
	Change Point	OOS	Monthly	0.007006253
C1	Control	Within	Daily	0.02242433
	Asymptote	Within	Daily	0.02247720
	Change Point	Within	Daily	**0.02212230**
C1	Control	Within	Monthly	0.007450628
	Asymptote	Within	Monthly	0.008678233
	Change Point	Within	Monthly	**0.007032987**
C1	Control	OOS	Daily	**0.02630727**
	Asymptote	OOS	Daily	0.02769865
	Change Point	OOS	Daily	0.02707440
C1	Control	OOS	Monthly	0.006037918
	Asymptote	OOS	Monthly	0.005115459
	Change Point	OOS	Monthly	**0.004834255**
C4	Control	Within	Daily	**0.02327867**
	Asymptote	Within	Daily	0.02413068
	Change Point	Within	Daily	**0.02327370**
C4	Control	Within	Monthly	**0.008220368**
	Asymptote	Within	Monthly	0.009892227
	Change Point	Within	Monthly	**0.008219834**
C4	Control	OOS	Daily	**0.02663727**
	Asymptote	OOS	Daily	0.02756272
	Change Point	OOS	Daily	**0.02676222**
C4	Control	OOS	Monthly	**0.011126124**
	Asymptote	OOS	Monthly	0.012295470
	Change Point	OOS	Monthly	0.012825395

head approaches the land surface. Specifically, this groundwater dependent ecosystem is capable of evapotranspiring at 75% of the potential limit as groundwater depth approaches the land surface.

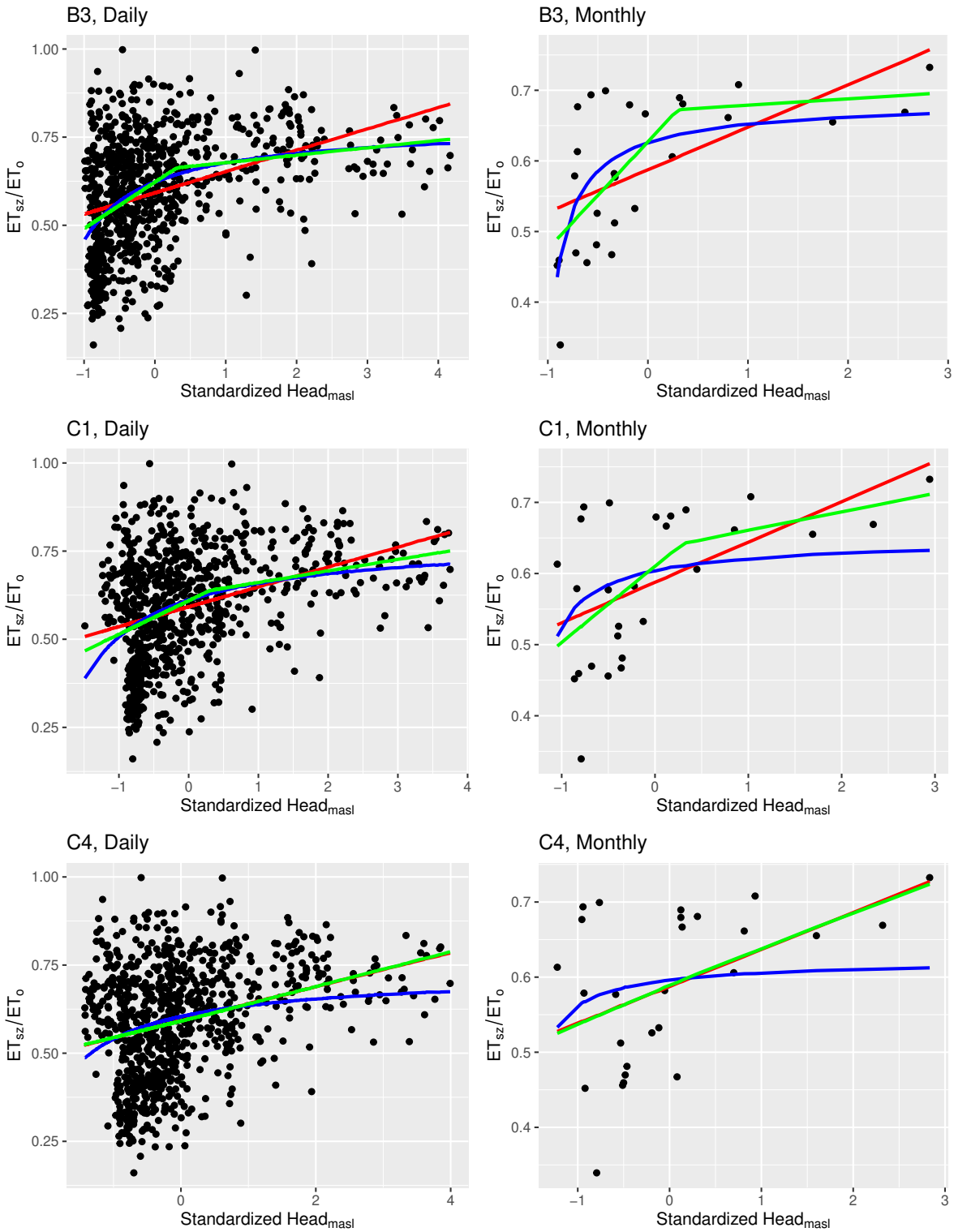

Figure 4.7: Daily (left) and monthly (right) data overlaid by the posterior mean prediction of y_i using the control (red), asymptote (blue), and change point (green) functions plotted as lines.

Although, this assertion deserves further testing using weather station data collected within other vegetation types. The c variable represents the depth of groundwater at which the growth increment of ET_{sz}/ET_o is zero. The mean value of c is 1.58 and the non-standardized value is 1256.675 masl which is the elevation of the water table required to instantiate growth of tamarisk next to the Arkansas River; though this variable is spatially dependent. The mean of the posterior distribution for the variance component is 0.021 which represents the error in estimating ET_{sz}/ET_o using an asymptotic function and groundwater head as the sole covariate.

There is a vast amount of literature that supports linear and change point models applied in groundwater modeling. However, a large amount of research has analyzed the groundwater-ET relationship using satellite imagery with a coarse temporal scale of analysis and thus makes it difficult to compare to this high-resolution study. Figure 4.9 provides the posterior distributions for the change point model parameters. The β_1, β_2, and cp parameters have ecohydrological interpretation. Both rates, β_1 and β_2, are positive for the shrub and phreatophyte PFG which indicates that as the water table approaches the land surface then evapotranspiration increases. Similar to the asymptote model, the slope of the line between evapotranspiration and groundwater head is steeper during the dormant season for the shrub and phreatophyte PFG. The slope of the line during the actively-growing season is positive but an order of magnitude smaller. The smaller slope during the actively-growing season indicates that increases in groundwater head (i.e., shallowing) result in smaller increases in evapotranspiration compared to the dormant season. The change point parameter marks the level of groundwater head at which tamarisk changes ecological state from dormant to actively growing. The posterior mean of the cp parameter is 0.3 and the non-standardized value is 1256.925 masl for the shrub PFG. The grass PFG, as shown in Figure 4.7, did not require a change point parameter to model the ET-groundwater process which is why β_2 overlaps zero in Figure 4.8 and Table 4.3 indicates and equal predictive score bewteen the change point and control models. The mean of the posterior distribution for the variance component is 0.021 which represents the error in estimating ET_{sz}/ET_o using a change point function and groundwater head as the sole covariate.

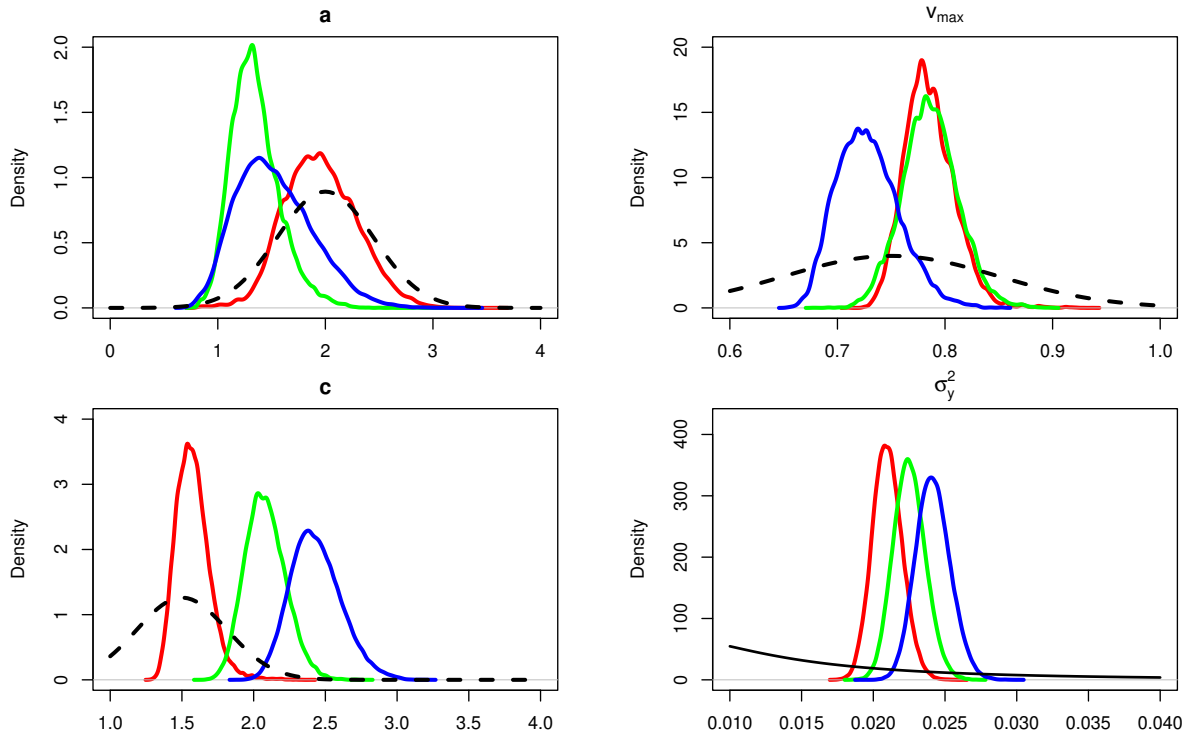

Figure 4.8: Marginal posterior distributions for the asymptote model applied to wells B3 (red), C1 (green), and C4 (blue). Priors for each parameter are shown dashed-black lines.

4.4 Conclusion

A Bayesian framework was used to derive and test which function form best fits the relationship between evapotranspiration and groundwater head at the point scale. The answer to the hypothesis is partially confirmed. The asymptote model is more predictive, based on MSPE scores, but it depends on PFG. The control model and the change point model had relatively equal MSPE scores for the grass PFG. The resulting MSPE scores of the asymptote model, with added model parameter interpretability, will be useful in future GDE research that compares GDE processes along a climatic gradient. Additionally, the asymptote function could be incorporated into a spatially distributed groundwater modeling scheme to test the performance against linear functions.

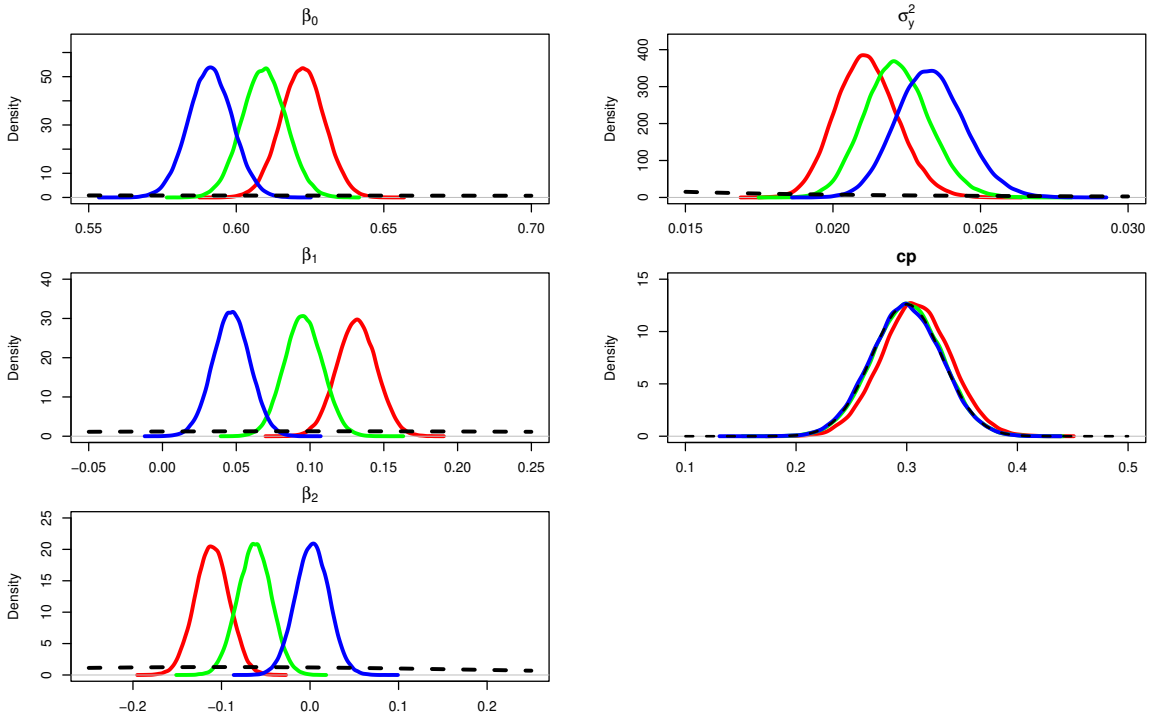

Figure 4.9: Marginal posterior distributions for the change point model applied to wells B3 (red), C1 (green), and C4 (blue). Priors for each parameter are shown solid black lines.

Chapter 5

General Conclusions

After the general introduction, the socio-hydrologic connections with a groundwater dependent ecosystem were examined using: a) output from a calibrated numerical model and a validated remote-sensing based evapotranspiration algorithm, b) hierarchical time series and spatial statistics, and c) a simple regression framework. Though each approach is independent, there is strong utility and shared characteristics among the three applications.

All three approaches examine vegetation patterns along the Arkansas River and quantify the connections between a riparian ecosystem and floodplain activity (e.g., agro-pastoral practices) in southeastern Colorado. In Chapter 2, a heavily calibrated numerical model is correlated with remote sensing estimates of riparian evapotranspiration to make inference on how return flows (e.g., groundwater depth), that have changed in timing and amount due to changing irrigation strategies, impact riparian water use. This initial investigation led to a complimentary analysis using spatial and temporal statistics and a further investigation into the groundwater-ET relationship. In Chapter 3, hierarchical time series and spatial statistics are used to connect NDVI, a proxy for riparian ET, to socio-hydrologic covariates while acknowledging autocorrelation. The time series and spatial statistics approach is easily repeatable and has application in data-rich and data-sparse watersheds. Chapter 4 is a simplistic and useful test of well-known theory in groundwater modeling: as groundwater depth decreases then evapotranspiration increases linearly. The Bayesian regression framework tests this theory given the implications for correctly modeling the groundwater-ET relationship in large-scale water balance studies that include GDEs as a component of the hydrosocial cycle.

5.1 Placement in Literature

Groundwater dependent ecosystems and linkages between an irrigated agricultural region and a natural ecosystem are well-studied topics but still require further investigation. The three research approaches presented here have advanced these topics and have relevance to current literature.

5.1.1 Catchment-Scale

The first approach involving MODFLOW and a remote sensing ET algorithm provided better understanding of groundwater depth thresholds and how ET changes during wet and dry years. Simulated estimates of groundwater depth within the riparian corridor agree well with studies in similar climates with similar vegetation type [19, 134, 135]. The relationship between groundwater depth and ET that changes slope beyond a given depth has relevance to deterministic functions used for groundwater-ET relationships in groundwater modeling schemes [37, 50, 115]. Remote sensing based estimates of riparian ET in Chapter 2 are higher than normally reported in literature for semi-arid climates [46], but the result regarding higher riparian ET in drier periods, owing to groundwater dependence to overcome water deficit, has been quantified in other riparian vegetation patches [9]. These results can be used to identify management focus areas where vegetation is sensitive to changes in groundwater levels. The following research approach in Chapter 3 complimented this analysis by examining exploratory covariates tied to agro-pastoral practices.

5.1.2 Space and Time Statistics

The second approach started with a time series analysis that derived trend, autocorrelation, and the strength of the connections between a riparian GDE and hydro-social variables. Temporal dependence, in the form of an AR(1) model, is accounted for in the time series to aid in understanding the impact of explanatory covariates [136]. Without the AR component, inference on the hydrological covariates would be optimistic [137]. At the chosen spatial scale of analysis, temporal trend was not identifiable at seven of eight subregions. Although, trend analysis in a similar climate with similar vegetation type found statistically significant downward trends of annual ri-

parian NDVI along the San Pedro River in Arizona [85]. Recent studies have used the Theil–Sen median trend analysis and the Mann–Kendall trend test [12, 24, 138]. The model I presented quantifies autoregressive behavior and monthly trend simultaneously and is easily repeatable in other data sparse and data rich regions. Zhu et al. (2021) analyzed the trend in gross primary production, temperature, precipitation and SPEI for eight vegetation regions in China. They used the Theil-Sen and the Mann-Kendall methods to quantify trend and removed lag-1 autocorrelation before using both methods on each time series. They noted that lag-1 serial correlation may overestimate the probability of detecting a significant trend [138]. Pei et al. (2019) analyzed NDVI time series from grassland sites in Inner Mongolia (CN) from 1982-2015 [139]. Additionally, in agreement with the time series approach in Chapter 3, Pei et al. (2019) found higher correlations between NDVI and climatic factors during within-growing season (April to October) analysis and concluded that NDVI was sensitive to climate factors with zero lag and with one-month lag but it depended on grassland type.

Spatial statistics can be used to analyze plant populations and make inference and hypotheses. Floodplain activity such as pastoral practices [140], cities and townships, and agriculture [24] have all been linked to riparian ecosystem changes. The conditional autoregressive model used in the spatial statistics part of the second approach addresses spatial correlation within each subregion while testing for linkages between floodplain activity and riparian vegetation density. Autocorrelation is included in the spatial statistics approach, much like the temporal approach, because accounting for autocorrelation (i.e., similarity among close locations) helps clarify the effect of different explanatory variables which improves inference [136, 141]. The conditional autoregressive model, like the one used in Chapter 3, is often used in vegetation prediction models used to quantify vegetation abundance across a landscape [136]. The impact of agriculture, or a disturbance, on riparian vegetation density has been reported in literature. Vanderhoof et al. (2019) documented decreases in peak-season riparian wetness, likely due to irrigation methodology change, in the upper Missouri headwaters [24]. Albano et al. (2020) found negative temporal trends in dryland riparian vegetation bordering low-intensity agriculture in Nevada [12]. These findings are used to

create hypotheses about the connection between floodplain activity and riparian ecosystems and the spatial statistics revealed that vegetation density was reduced in riparian corridors that neighbor permanent irrigated land dry-up. Like in the spatial statistics approach in Chapter 3, the usefulness of categorical predictor variables is common in literature. Harner and Stanford (2003) used multiple regression with indicator variables to determine the effects of gaining and losing streams on cottonwood growth and age in Montana [142]. These results can be used to sustain GDEs in Colorado and other western states that have similar ecosystems under anthropogenic pressures.

5.1.3 Point-Scale

The third approach that uses a regression methodology on the groundwater-ET relationship has many linkages to research in GDEs and groundwater modeling. Based on the simulated findings from Chapter 2, further investigating the groundwater-ET relationship with point-scale observations was of interest. In general, the findings from Chapter 4 suggest that the process between groundwater and ET is curvi-linear but depends on plant functional group. This finding has been observed in other point-scale studies of riparian vegetation in Nebraska on a daily basis [13] and in herbaceous wetlands in an arid basin of Colorado on an annual basis [125]. Asymptotic functions are not commonly used in ecohydrology but several studies have reported a linear change point [71] and linear relationship [7, 14] between groundwater and ET. Although, asymptotic-like relationships have been found between NDVI and groundwater in a semi-arid region of China but depended on vegetation type [7]. The findings that support the use of linear and change point linear models have relevance in groundwater modeling studies [37, 115].

5.2 Placement in Broader Context

More broadly, this research fits in with recent literature surrounding critical zone areas that are connected to Endangered Rivers [2]. Groundwater-dependent ecosystems can provide numerous ecosystem services and mark the presence of water resources. Timing is tantamount in sustaining these ecosystems in the face of resource demand and climate change. Advancement of computing

capabilities, sensor networks, and remote-sensing platforms are creating an exciting future for GDE assessment and monitoring. Going forward, a scale-enlightened synergistic application of mass-based (i.e., numerical models) and energy-based remote sensing applications will be critical to sustaining groundwater dependent ecosystems.

Bibliography

[1] F. A. Ward and M. Pulido-Velazquez. Water conservation in irrigation can increase water use. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47):18215–18220, 2008.

[2] A. E. Petri. Top 10 Most Endangered Rivers in the U.S. *National Geographic*, 2017.

[3] T. M. Doody, O. Barron, K. Dowsley, I. Emelyanova, J. Fawcett, I. Overton, J. Pritchard, A. Van Dijk, and G. Warren. Continental mapping of groundwater dependent ecosystems: A methodological framework to integrate diverse data and expert opinion. *Journal of Hydrology: Regional Studies*, 10:61–81, 2017.

[4] S. Jasechko, H. Seybold, D. Perrone, Y. Fan, and J. Kirchner. Widespread potential loss of streamflow into underlying aquifers across the USA. *Nature*, 591(7850):391–395, 2021.

[5] S. P. Loheide and S. Gorelick. A local-scale, high-resolution evapotranspiration mapping algorithm (ETMA) with hydroecological applications at riparian meadow. *Remote Sensing of Environment*, 98(2-3):182–200, 2005.

[6] B. R. Murray, M. J. B. Zeppel, G. C. Hose, and D. Eamus. Groundwater-dependent ecosystems in Australia: It's more than just water for rivers. *Ecological Management and Restoration*, 4:53–59, 2003.

[7] J. Lv, X. Wang, Y. Zhou, K. Qian, L. Wan, D. Eamus, and Z. Tao. Groundwater-dependent distribution of vegetation in hailiutu river catchment, a semi-arid region in china. *Ecohydrology*, 1:142–149, 2012.

[8] D. Eamus and R. Froend. Groundwater-dependent ecosystems: the where, what and why of GDEs. *Australian Journal of Botany*, 54:91–96, 2006.

[9] B. Sommer, D. Boggs, G. Boggs, A. van Dijk, and R. Froend. Spatio-temporal patterns of evapotranspiration from groundwater-dependent vegetation. *Ecohydrology*, 9(8):1620–1629, 2016.

[10] B. Kløve, P. Ala-aho, G. Bertrand, Z. Boukalova, A. Ertürk, N. Goldscheider, J. Ilmonen, N. Karakaya, H. Kupfersberger, J. Kvœrner, A. Lundberg, M. Mileusnić, A. Moszczynska, T. Muotka, E. Preda, P. Rossi, D. Siergieiev, J. Šimek, P. Wachniew, V. Angheluta, and A. Widerlund. Groundwater dependent ecosystems. Part I: Hydroecological status and trends. *Environmental Science and Policy*, 14(7):770–781, 2011.

[11] O. Johansen, D. Andersen, R. Ejrnæs, and M. Pedersen. Relations between vegetation and water level in groundwater dependent terrestrial ecosystems (GWDTEs). *Limnologica*, 68:130–141, 2018.

[12] C. Albano, K. Mcgwire, M. Hausner, D. Mcevoy, C. Morton, and J. Huntington. Drought sensitivity and trends of riparian vegetation vigor in nevada, usa (1985–2018). *Remote Sensing*, 12(1362), 2020.

[13] W. Yue, T. Wang, T. E. Franz, and X. Chen. Spatiotemporal patterns of water table fluctuations and evapotranspiration induced by riparian vegetation in a semiarid area. *Water Resources Research*, 3:1948–1960, 2016.

[14] J. Missik, H. Liu, Z. Gao, M. Huang, X. Chen, E. Arntzen, D. Mcfarland, and B. Verbeke. Groundwater regulates interannual variations in evapotranspiration in a riparian semiarid ecosystem. *Journal of Geophysical Research: Atmospheres*, 126, 2021.

[15] R. Scott, W. Cable, T. Huxman, P. Nagler, M. Hernandez, and D. Goodrich. Multiyear riparian evapotranspiration and groundwater use for a semiarid watershed. *Journal of Arid Environments*, 72(7):1232–1246, 2008.

[16] R. G. Allen, M. Tasumi, A. Morse, R. Trezza, J. L. Wright, W. Bastiaanssen, W. Kramber, I. Lorite, and C. W. Robison. Satellite-Based Energy Balance for Mapping Evapotranspi-

ration with Internalized Calibration (METRIC)—Applications. *Journal of Irrigation and Drainage Engineering*, 4:395–406, 2007.

[17] E. P. Glenn, P. L. Nagler, and A. R. Huete. Vegetation index methods for estimating evapotranspiration by remote sensing. *Surveys in Geophysics*, 31:531–555, 2010.

[18] M. McCabe, B. Aragon, R. Houborg, and J. Mascaro. Cubesats in hydrology: Ultrahigh-resolution insights into vegetation dynamics and terrestrial evaporation. *Water Resources Research*, 53(12):10017–10024, 2017.

[19] F. Orellana, P. Verma, S. P. Loheide, and E. Daly. Monitoring and modeling water-vegetation interactions in groundwater-dependent ecosystems. *Reviews of Geophysics*, 3:501–512, 2012.

[20] M. Hausner, J. Huntington, C. Nash, C. Morton, D. McEvoy, D. Pilliod, K. Hegewisch, B. Daudert, J. Abatzoglou, and G. Grant. Assessing the effectiveness of riparian restoration projects using landsat and precipitation data from the cloud-computing application climateengine.org. *Ecological Engineering*, 120:432–440, 2018.

[21] K. Anderson, D. Fawcett, A. Cugulliere, S. Benford, D. Jones, and R. Leng. Vegetation expansion in the subnival hindu kush himalaya. *Global Change Biology*, (14919):1–18, 2019.

[22] E. Kendy and J. D. Bredehoeft. Transient effects of groundwater pumping and surface-water-irrigation returns on streamflow. *Water Resources Research*, 8, 2006.

[23] B. R. Scanlon, R. W. Healy, and P. G. Cook. Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, 10:18–39, 2002.

[24] M. Vanderhoof, J. Christensen, and L. Alexander. Influence of multi-decadal land use, irrigation practices and climate on riparian corridors across the Upper Missouri River Headwaters Basin, Montana. *Hydrology and Earth System Sciences*, 23:4269–4292, 2019.

[25] B. Osborn. Lower Arkansas River Watershed Plan: John Martin Reservoir to Stateline. Technical report, Colorado Water Institute - Colorado State University, 2019.

[26] J. D. Niemann, B. M. Lehman, T. K. Gates, N. U. Hallberg, and A. Elhaddad. Impact of Shallow Groundwater on Evapotranspiration Losses from Uncultivated Land in an Irrigated River Valley. *Journal of Irrigation and Drainage Engineering*, 8:501–512, 2011.

[27] Core Writing Team, H. Lee, and J. (eds.) Romero. Summary for policymakers. in: Climate change 2023: Synthesis report. a report of the intergovernmental panel on climate change. Technical report, Intergovernmental Panel on Climate Change, IPCC, Geneva, Switzerland, March 2023.

[28] D. Eamus, S. Zolfaghar, R. Villalobos-Vega, J. Cleverly, and A. Huete. Groundwater-dependent ecosystems: Recent insights from satellite and field-based studies. *Hydrology and Earth System Sciences*, 10:4229–4256, 2015.

[29] L. Foglia, A. McNally, and T. Harter. Coupling a spatiotemporally distributed soil water budget with stream-depletion functions to inform stakeholder-driven management of groundwater-dependent ecosystems. *Water Resources Research*, pages 7292–7310, 2013.

[30] G. Blöschl and M. Sivapalan. Scale Issues In Hydrological Modelling: A Review. *Hydrological Processes*, 9:251–290, 1995.

[31] J. Linton and J. Budds. The hydrosocial cycle: Defining and mobilizing a relational-dialectical approach to water. *Geoforum*, pages 170–180, 2014.

[32] G. N. Zaimes, R. C. Schultz, and T. M. Isenhart. Streambank soil and phosphorus losses under different riparian land-uses in Iowa. *Journal of the American Water Resources Association*, 4:935–947, 2008.

[33] T. M. Doody, M. J. Colloff, M. Davies, V. Koul, R. G. Benyon, and P. L. Nagler. Quantifying water requirements of riparian river red gum (Eucalyptus camaldulensis) in the Murray-

Darling Basin, Australia - implications for the management of environmental flows. *Ecohydrology*, 8:1471–1487, 2015.

[34] R. T. Bailey, T. K. Gates, and E. C. Romero. Assessing the effectiveness of land and water management practices on nonpoint source nitrate levels in an alluvial stream-aquifer system. *Journal of Contaminant Hydrology*, 2015.

[35] C. D. Shultz, R. T. Bailey, T. K. Gates, B. E. Heesemann, and E. D. Morway. Simulating selenium and nitrogen fate and transport in coupled stream-aquifer systems of irrigated regions. *Journal of Hydrology*, pages 512–529, 2018a.

[36] R. Doble, C. Simmons, I. Jolly, and G. Walker. Spatial relationships between vegetation cover and irrigation-induced groundwater discharge on a semi-arid floodplain, Australia. *Journal of Hydrology*, 1-2:75–97, 2006.

[37] K. J. Baird and T. Maddock. Simulating riparian evapotranspiration: A new methodology and application for groundwater models. *Journal of Hydrology*, 1-4:176–190, 2005.

[38] W. G. M. Bastiaanssen, H. Pelgrum, J. Wang, Y. Ma, J. F. Moreno, G. J. Roerink, and T. van der Wal. A remote sensing surface energy balance algorithm for land (SEBAL), Part 1: Formulation. *Journal of Hydrology*, pages 198–212, 1998.

[39] A. Elhaddad and L. A. Garcia. Surface Energy Balance-Based Model for Estimating Evapotranspiration Taking into Account Spatial Variability in Weather. *Journal of Irrigation and Drainage Engineering*, 6:681–689, 2008.

[40] J. H. Chen, C. E. Kan, C. H. Tan, and S. F. Shih. Use of spectral information for wetland evapotranspiration assessment. *Agricultural Water Management*, 3:239–248, 2002.

[41] L. Jiang and S. Islam. An intercomparison of regional latent heat flux estimation using remote sensing data. *International Journal of Remote Sensing*, 11:2221–2236, 2003.

[42] P. L. Nagler, J. Cleverly, E. Glenn, D. Lampkin, A. Huete, and Z. Wan. Predicting riparian evapotranspiration from MODIS vegetation indices and meteorological data. *Remote Sensing of Environment*, 1:17–30, 2005.

[43] Z. Samani, A. S. Bawazir, R. K. Skaggs, M. P. Bleiweiss, A. Piñon, and V. Tran. Water Use by Agricultural Crops and Riparian Vegetation: An Application of Remote Sensing Technology. *Journal of Contemporary Water Research Education*, 1:8–13, 2009.

[44] G. B. Senay, S. Bohms, R. K. Singh, P. H. Gowda, N. M. Velpuri, H. Alemu, and J. P. Verdin. Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach. *Journal of the American Water Resources Association*, 3:577–591, 2013.

[45] G. Petropoulos, T. N. Carlson, M. J. Wooster, and S. Islam. A review of Ts/VI remote sensing based methods for the retrieval of land surface energy fluxes and soil surface moisture. *Progress in Physical Geography*, 2:224–250, 2009.

[46] G. B. Senay, M. Friedrichs, R. K. Singh, and N. M. Velpuri. Evaluating Landsat 8 evapotranspiration for water use mapping in the Colorado River Basin. *Remote Sensing of Environment*, pages 171–185, 2016.

[47] G. B. Senay, M. Schauer, M. Friedrichs, N. M. Velpuri, and R. K. Singh. Satellite-based water use dynamics using historical Landsat data (1984–2014) in the southwestern United States. *Remote Sensing of Environment*, pages 98–112, 2017.

[48] E. P. Glenn, P. L. Nagler, K. Morino, and K. R. Hultine. Phreatophytes under stress: Transpiration and stomatal conductance of saltcedar (Tamarix spp.) in a high-salinity environment. *Plant and Soil*, 1-2:655–672, 2013.

[49] K. Khand, S. Taghvaeian, and L. Hassan-Esfahani. Mapping annual riparian water use based on the single-satellite-scene approach. *Remote Sensing*, 8, 2017.

[50] F. D. Tillman, S. M. Wiele, and D. R. Pool. A comparison of estimates of basin-scale soil-moisture evapotranspiration and estimates of riparian groundwater evapotranspiration with implications for water budgets in the Verde Valley, Central Arizona, USA. *Journal of Arid Environments*, pages 278–291, 2016.

[51] E. P. Glenn, R. L. Scott, U. Nguyen, and P. L. Nagler. Wide-area ratios of evapotranspiration to precipitation in monsoon-dependent semiarid vegetation communities. *Journal of Arid Environments*, pages 84–95, 2015.

[52] E. D. Morway, T. K. Gates, and R. G. Niswonger. Appraising options to reduce shallow groundwater tables and enhance flow conditions over regional scales in an irrigated alluvial aquifer system. *Journal of Hydrology*, pages 216–237, 2013.

[53] B. Dewandel, J. M. Gandolfi, D. de Condappa, and S. Ahmed. An efficient methodology for estimating irrigation return flow coefficients of irrigated crops at watershed and seasonal scale. *Hydrological Processes*, 11:1700–1712, 2008.

[54] T. K. Gates, J. T. Cox, and K. H. Morse. Uncertainty in mass-balance estimates of regional irrigation-induced return flows and pollutant loads to a river. *Journal of Hydrology: Regional Studies*, pages 193–210, 2018.

[55] R. T. Bailey, T. C. Wible, M. Arabi, R. M. Records, and J. Ditty. Assessing regional-scale spatio-temporal patterns of groundwater–surface water interactions using a coupled SWAT-MODFLOW model. *Hydrological Processes*, 23:4420–4433, 2016.

[56] R. G. Bailey. Identifying ecoregion boundaries. *Environmental Management*, pages 14–26, 2004.

[57] T. K. Gates, L. A. Garcia, R. A. Hemphill, E. D. Morway, and A. Elhaddad. Irrigation practices, water consumption, and return flows in Colorado's Lower Arkansas River Basin: Field and model investigations. *CWI Completion Report No. 221 CAES Report No. TR12-10*, 2012.

[58] D. M. Merritt, M. E. Manning, and N. Hough-Snee. The National Riparian Core Protocol: A riparian vegetation monitoring protocol for wadeable streams of the conterminous United States Gen. Tech. Rep. RMRS-GTR-367. Technical report, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO 80523, 2017.

[59] G. Kittel, E. Vanwie, M. Damm, R. Rondeau, S. Kettler, A. Mcmullen, and J. Sanderson. A Classification of Riparian Wetland Plant Associations of Colorado: User Guide to the Classification Project. Technical report, Colorado Natural Heritage Program, Colorado State University, Fort Collins, CO 80523, September 1999.

[60] R. G. Allen, I. A. Walter, R. Elliott, T. Howell, D. Itenfisu, and M. Jensen. The ASCE Standardized Reference Evapotranspiration Equation. 2005.

[61] H. S. Al Wahaibi. Evaluating The ASCE Standardized Penman-Monteith Equation and Developing Crop Coefficients of Alfalfa Using A Weighing Lysimeter in Southeast Colorado. Master's thesis, Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, 2011.

[62] A. A. Andales, D. Straw, T. H. Marek, L. H. Simmons, M. E. Bartolo, and T. W. Ley. Design and Construction of a Precision Weighing Lysimeter in Southeast Colorado. *Transactions of the ASABE*, 2:509–521, 2018.

[63] A. Subedi, J. L. Chávez, and A. A. Andales. ASCE-EWRI standardized Penman-Monteith evapotranspiration (ET) equation performance in southeastern Colorado. *Agricultural Water Management*, pages 74–80, 2017.

[64] G. B. Senay, M. Budde, J. P. Verdin, and A. M. Melesse. A coupled remote sensing and simplified surface energy balance approach to estimate actual evapotranspiration from irrigated fields. *Sensors*, 6:979–1000, 2007.

[65] J. Kim and T. S. Hogue. Evaluation of a MODIS triangle-based evapotranspiration algorithm for semi-arid regions. *Journal of Applied Remote Sensing*, 1, 2013.

[66] Z. Zhu and C. E. Woodcock. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, pages 5009–5021, 2012.

[67] M. Hollander and D. A. Wolfe. *Nonparametric Statistical Methods*, volume 3. New York, NY: John Wiley and Sons, 1973.

[68] J. P. Royston. Algorithm AS 181: The W Test for Normality. *Applied Statistics*, 2:176–180, 1982.

[69] G. N. Zaimes, R. C. Schultz, and T. M. Isenhart. Significance Testing of the Spearman Rank Correlation Coefficient. *Journal of the American Statistical Association*, 339:578–580, 1972.

[70] R. F. Bay and A. A. Sher. Success of active revegetation after Tamarix removal in riparian ecosystems of the Southwestern United States: A quantitative assessment of past restoration projects. *Restoration Ecology*, 1:113–128, 2008.

[71] R. C. Doble and R. S. Crosbie. Review: Current and emerging methods for catchment-scale modelling of recharge and evapotranspiration from shallow groundwater. *Hydrogeology Journal*, 1:3–23, 2017.

[72] C. D. Shultz, T. K. Gates, and R. T. Bailey. Evaluating best management practices to lower selenium and nitrate in groundwater and streams in an irrigated river valley using a calibrated fate and reactive transport model. *Journal of Hydrology*, pages 299–312, 2018b.

[73] S. P. Loheide and E. G. Booth. Effects of changing channel morphology on vegetation, groundwater, and soil moisture regimes in groundwater-dependent ecosystems. *Geomorphology*, 126:364–376, 2011.

[74] M. Ronayne, J. Roudebush, and J. Stednick. Analysis of managed aquifer recharge for retiming streamflow in an alluvial river. *Journal of Hydrology*, 544:373–382, 2017.

[75] T. K. Gates, G. H. Steed, J. D. Niemann, and J. W. Labadie. Data for improved water management in Colorado's Arkansas River Basin: Hydrological and water quality studies. 2016.

[76] W. Berghuijs, E. Luijendijk, C. Moeck, Y. van der Velde, and S. Allen. Global recharge data set indicates strengthened groundwater connection to surface fluxes. *Geophysical Research Letters*, 49(23), 2022.

[77] K. R. Hultine, R. Froend, D. Blasini, S. E. Bush, M. Karlinski, and D. F. Koepke. Hydraulic traits that buffer deep-rooted plants from changes in hydrology and climate. *Hydrological Processes*, 34(2):209–222, 2020.

[78] J. Yao, H. Liu, J. Huang, Z. Gao, G. Wang, D. Li, H. Yu, and X. Chen. Accelerated dryland expansion regulates future variability in dryland gross primary production. *Nature Communications*, 20(11):1665, 2020.

[79] M. B. Hooten and T. J. Hefley. *Bringing Bayesian Models to Life*. Chapman and Hall/CRC, 2019.

[80] B. Jiang, S. Liang, J. Wang, and Z. Xiao. Modeling MODIS LAI time series using three statistical methods. *Remote Sensing of Environment*, 14:1432–1444, 2010.

[81] A. Montanari, R. Rosso, and M. Taqqu. Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resources Research*, 33(5):1035–1044, 1997.

[82] R. L. Bras and I. Rodriguez-Iturbe. *Random Functions and Hydrology*. Addison-Wesley Publ. Co., Massachusetts, 1985. ISBN:0-486-67626-9.

[83] J. D. Salas, G. Q. Tabios III, and P. Bartolini. Approaches to multivariate modeling of water resources time series. *Journal of American Water Resources Association*, 21(4):683–708, 1985.

[84] R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications - With R Examples*. Springer, 4th ed, 2016.

[85] U. Nguyen, E. Glenn, P. Nagler, and R. Scott. Long-term decrease in satellite vegetation indices in response to environmental variables in an iconic desert riparian ecosystem: the Upper San Pedro, Arizona, United States. *Ecohydrology*, 8(4):610–625, 2015.

[86] R. de Jong, S. de Bruin, A. de Wit, M. E. Schaepman, and D. L. Dent. Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sensing of Environment*, 115:692–702, 2010.

[87] Y. Lin, X. Xin, H. Zhang, and X. Wang. The implications of serial correlation and time-lag effects for the impact study of climate change on vegetation dynamics – a case study with Hulunber meadow steppe, Inner Mongolia. *International Journal of Remote Sensing*, 36(19-20):5031–5044, 2015.

[88] T. Hefley, K. Broms, B. Brost, F. Buderman, S. Kay, H. Scharf, J. Tipton, P. Williams, and M. Hooten. The basis function approach for modeling autocorrelation in ecological data. *Ecology: Concepts and Synthesis*, 98(3):632–646, 2016.

[89] N. T. Hobbs and M. B. Hooten. *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press, 2015.

[90] M. Heaton, A. Datta, A. Finley, R. Furrer, J. Guinness, R. Guhaniyogi, F. Gerber, R. Gramacy, D. Hammerling, M. Katzfuss, F. Lindgren, D. Nychka, F. Sun, and A. Zammit-Mangion. A case study competition among methods for analyzing large spatial data. *Journal of Agricultural, Biological, and Environmental Statistics*, 24(3):398–425, 2019.

[91] J. Huntington, K. McGwire, C. Morton, K. Snyder, S. Peterson, T. Erickson, R. Niswonger, R. Carroll, G. Smith, and R. Allen. Assessing the role of climate and resource management on groundwater dependent ecosystem changes in arid environments with the landsat archive. *Remote Sensing of Environment*, 185:186–197, 2016.

[92] J. A. Ryan, J. M. Ulrich, R. Bennett, and C. Joy. xts: extensible time series, 2020.

[93] N. Pettorelli, J. Vik, A. Mysterud, J. Gaillard, C. Tucker, and N. Stenseth. Using the satellite-derived NDVI to assess ecological responses to environmental change. *TRENDS in Ecology and Evolution*, 20(9):503–510, 2005.

[94] Case US Supreme Court Case 514. Amended appendix c.1, hydrologic - institutional model: Model documentation, Amended September 2011.

[95] A. Holm, S. Cridland, and M. Roderick. The use of time-integrated NOAA NDVI data and rainfall to assess landscape degradation in the arid shrubland of Western Australia. *Remote Sensing of Environment*, 85(2):145–158, 2003.

[96] B. C. Reed, T. R. Loveland, and L. L. Tiezen. An approach to using AVHRR data to monitor U.S. great plains grasslands. *Geocarto International*, 11(3):13–22, 1996.

[97] J. T. Abatzoglou, D. J. McEvoy, and K. T. Redmond. The west wide drought tracker: Drought monitoring at fine spatial scales. *Bulletin of the American Meteorological Society*, 98(9), 2017.

[98] D. P. Bunting, S. A. Kurca, E. P. Glenn, P. L. Nagler, and R. L. Scott. Insights for empirically modeling evapotranspiration influenced by riparian and upland vegetation in semiarid regions. *Journal of Arid Environments*, 111:42–52, 2014.

[99] U.S. Geological Survey. National water information system data available on the world wide web (usgs water data for the nation), 2016.

[100] E. Paradis. ape: Analyses of phylogenetics and evolution, 2020.

[101] P. Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37:17–23, 1950.

[102] K. Johansen, N. C. Coops, S. E. Gergel, and Y. Stange. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment*, 110:29–44, 2007.

[103] P. S. Eagleson. Climate, soil and vegetation. Part II: The distribution of annual precipitation derived from observed storm sequences. *Water Resources Research*, 14(5):713–721, 1978.

[104] J. A. Ramirez and S. U. S. Senarath. A Statistical-Dynamical Parameterization of Canopy Interception and Land Surface-Atmosphere Interactions. *Journal of Climate*, 13:4050–4063, 2000.

[105] X. Chen, Y. Rubin, S. Ma, and D. Baldocchi. Observations and stochastic modeling of soil moisture control on evapotranspiration in a californian oak savanna. *Water Resource Research*, 44(W08409), 2008.

[106] P. Bartolini, J. Salas, and J. Obeysekera. Multivariate Periodic ARMA(1, 1) Processes. *Water Resources Research*, 24(8):1237–1246, 1988.

[107] M. Parsons and M. C. Thoms. Patterns of vegetation greenness during flood, rain and dry resource states in a large, unconfined floodplain landscape. *Journal of Arid Environments*, 88:24–38, 2013.

[108] P. Nagler, T. Doody, E. Glenn, C. Jarchow, A. Barreto-Muñoz, and K. Didan. Wide-area estimates of evapotranspiration by red gum (Eucalyptus camaldulensis) and associated vegetation in the Murray-Darling River Basin, Australia. *Hydrological Processes*, 30(9):1376–1387, 2016.

[109] N. M. Velpuri, G. B. Senay, M. C. Schauer, A. Garcia, R. K. Singh, M. Friedrichs, S. Kagone, J. Haynes, and T. Conlon. Evaluation of hydrologic impact of an irrigation curtailment program using landsat satellite data. *Hydrological Processes*, 34(8):1697–1713, 2020.

[110] N. Sims and M. Colloff. Remote sensing of vegetation responses to flooding of a semi-arid floodplain: Implications for monitoring ecological effects of environmental flows. *Ecological Indicators*, 18:387–391, 2012.

[111] K. P. Overmars, G. H. J. de Koning, and A. Veldkamp. Spatial autocorrelation in multi-scale land use models. *Ecological Modelling*, 164:257–270, 2003.

[112] M. B. Hooten and C. K. Wikle. Shifts in the spatio-temporal growth dynamics of shortleaf pine. *Environmental and Ecological Statistics*, 14:207–227, 2007.

[113] A. Comber and M. Wulder. Considering spatiotemporal processes in big data analysis: Insights from remote sensing of land cover and land use. *Transactions in GIS*, 23:879–891, 2019.

[114] A. W. Harbaugh. Modflow-2005, the u.s. geological survey modular ground-water model – the ground-water flow process. *U.S. Geological Survey Techniques and Methods 6-A16*, pages 1–253, 2005.

[115] E. R. Banta. The u.s. geological survey modular ground-water model—documentation of packages for simulating evapotranspiration with a segmented function (ets1) and drains with return flow (drt1). *U.S. Geological Survey Open File*, page 127 pp., 2000.

[116] N. Shah, M. Nachabe, and M. Ross. Extinction depth and evapotranspiration from ground water under selected land covers. *Ground Water*, 45(3):329–338, 2007.

[117] W. D. Nichols. Regional groundwater evapotranspiration and groundwater budgets, great basin nevada. *U.S. Geol. Surv. Prof. Pap. A1–A13*, 2000.

[118] D. Groeneveld. Remotely-sensed groundwater evapotranspiration from alkali scrub affected by declining water table. *Journal of Hydrology*, 358(3-4):294–303, 2008.

[119] W. D. Nichols. Groundwater discharge by phreatophyte shrubs in the great basin as related to depth to groundwater. *Water Resources Research*, 30:3265–3274, 1994.

[120] S. Tamea, F. Laio, L. Ridolfi, P. D'Odorico, and I. Rodriguez-Iturbe. Ecohydrology of groundwater-dependent ecosystems: 2. stochastic soil moisture dynamics. *Water Resources Research*, 45(5), 2009.

[121] S. Tamea, R. Muneepeerakul, F. Laio, L. Ridolfi, and I. Rodriguez-Iturbe. Stochastic description of water table fluctuations in wetlands. *Geophysical Research Letters*, 37(6), 2010.

[122] F. Laio, S. Tamea, L. Ridolfi, P. D'Odorico, and I. Rodriguez-Iturbe. Ecohydrology of groundwater-dependent ecosystems: 1. stochastic water table dynamics. *Water Resources Research*, 45(5), 2009.

[123] T. Maddock, K. Baird, R. Hanson, W. Schmid, and H. Ajami. Rip-et: A riparian evapotranspiration package for modflow-2005. *U.S. Geological Survey Techniques and Methods 6-A39*, pages 1–76, 2012.

[124] R. G. Allen, L. Pereira, D. Raes, and M. Smith. Crop evapotranspiration: Guidelines for computing crop water requirements. *FAO Irrigation and Drainage Paper 56. Rome, Italy: Food and Agriculture Organisation*, 1998.

[125] J. Sanderson and D. Cooper. Ground water discharge by evapotranspiration in wetlands of an arid intermountain basin. *Journal of Hydrology*, 351(3-4):344–359, 2008.

[126] I. Rodriguez-Iturbe, D. Entekhabi, and R. L. Bras. Nonlinear dynamics of soil moisture at climate scales 1. stochastic analysis. *Water Resources Research*, 27:1899–1906, 1991.

[127] J. A. Ramirez, M. T. Hobbins, and T. C. Brown. Observational evidence of the complementary relationship in regional evaporation lends strong support for bouchet's hypothesis. *Geophysical Research Letters*, 32(L15401), 2005.

[128] P. D'Odorico, F. Laio, and L. Ridolfi. Noise-induced stability in dryland plant ecosystems. *Proc. Natl. Acad. Sci.*, 102:10819–10822, 2005.

[129] I. Rodriguez-Iturbe, P. D'Odorico, F. Laio, L. Ridolfi, and S. Tamea. Challenges in humid land ecohydrology: Interactions of water table and unsaturated zone with climate, soil, and vegetation. *Water Resources Research*, 43, 2007.

[130] L. Ridolfi, P. D'Odorico, and F. Laio. Effect of vegetation–water table feedbacks on the stability and resilience of plant ecosystems. *Water Resources Research*, W01201, 2006.

[131] A. Huizenga, R. Bailey, and T. Gates. Stream-aquifer and in-stream processes affecting nitrogen along a major river and contributing tributary. *Journal of Contaminant Hydrology*, 199:24–35, 2017.

[132] A. Gelman and J. Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, 2009.

[133] K. Johnson and R. Goody. The original michaelis constant: Translation of the 1913 michaelis-menten paper. *Biochemistry*, 50(39):8264–8269, 2011.

[134] D. Cooper, J. Sanderson, D. Stannard, and D. Groeneveld. Effects of long-term water table drawdown on evapo- transpiration and vegetation in an arid region phreatophyte community. *Journal of Hydrology*, 325(1-4):21–34, 2006.

[135] D. Busch, N. Ingraham, and S. Smith. Water uptake in woody riparian phreatophytes of the southwestern united states: A stable isotope study. *Ecological Applications*, 2(4):450–459, 1992.

[136] R. Miller and N. Pinter. Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*, 202(3-4):225–242, 2007.

[137] S. Yue, P. Pilon, B. Phinney, and G. Cavadias. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16(9):1807–1829, 2002.

[138] X. Zhu, S. Zhang, T. Liu, and Y. Liu. Impacts of heat and drought on gross primary productivity in china. *Remote Sensing*, 13(3):1–22, 2021.

[139] Z. Pei, S. Fang, W. Yang, L. Wang, M. Wu, Q. Zhang, W. Han, and D. Khoi. The relationship between ndvi and climate factors at different monthly time scales: A case study of grasslands in inner mongolia, china (1982–2015). *Sustainability*, 11(7243), 2019.

[140] A. Jansen and A. Robertson. Relationships between livestock management and the ecological condition of riparian habitats along an australian floodplain river. *Journal of Applied Ecology*, 38(1):63–75, 2001.

[141] F. Dormann, M. McPherson, Araújo B., R. Bivand, J. Bolliger, G. Carl, G. Davies, A. Hirzel, W. Jetz, W. Kissling, I. Kühn, R. Ohlemüller, R. Peres-Neto, B. Reineking, B. Schröder, M. Schurr, and R. Wilson. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, 30(5):609–628, 2007.

[142] M. Harner and J. Stanford. Differences in cottonwood growth between a losing and gaining reach of an alluvial floodplain. *Ecology*, 84(6):1453–1458, 2003.

Appendix A

Chapter 3: Supplementary Material

A.1 Model Posteriors

A.1.1 Posterior for Model 1

$$[\boldsymbol{\beta}, \sigma^2 | \mathrm{y}_t] \propto \left(\prod_{t=1}^{T} [\mathrm{y}_t | \boldsymbol{\beta}, \sigma^2] \right) [\sigma^2][\boldsymbol{\beta}] \tag{A.1}$$

A.1.2 Posterior for Model 2

$$[\boldsymbol{\beta}, \alpha, \sigma^2_{\mathrm{y}}, \sigma^2_{\mathrm{z}}, \mathrm{z}_t | \mathrm{y}_t] \propto \left(\prod_{t=1}^{T} [\mathrm{y}_t | \boldsymbol{\beta}, \mathrm{z}_t, \sigma^2_y][\mathrm{z}_t | \alpha, \mathrm{z}_{t-1}, \sigma^2_{\mathrm{z}}] \right) [\sigma^2_z][\sigma^2_y][\boldsymbol{\beta}][\mathrm{z}_0][\alpha] \tag{A.2}$$

A.1.3 Posterior for Model 3

$$[\rho, \boldsymbol{\beta}, \boldsymbol{\eta}_d, \sigma^2_{\mathrm{d}}, \sigma^2_{\eta_d} | \mathbf{y}_d] \propto \left(\prod_{i=1}^{N_d} [\mathrm{y}_d | \boldsymbol{\beta}, \boldsymbol{\eta}_d, \sigma^2_{\mathrm{d}}][\boldsymbol{\eta}_d | \rho, \sigma^2_{\eta_d}] \right) [\sigma^2_{\eta_d}][\sigma^2_{\mathrm{d}}][\boldsymbol{\beta}][\boldsymbol{\eta}_d][\rho] \tag{A.3}$$

A.2 Full Conditional Statements

A.2.1 Model 1

For the vector of regression coefficients $\boldsymbol{\beta}$:

$$[\boldsymbol{\beta} | \cdot] \propto \left(\prod_{t=1}^{T} [\mathrm{y}_t | \mathbf{x}'_t \boldsymbol{\beta}, \sigma^2] \right) [\boldsymbol{\beta}] \tag{A.4}$$

Complete the square and combine like terms to get a multivariate normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \sum \left(\mathrm{y}'_t (\sigma^2)^{-1} \mathbf{x}_t \right) + \boldsymbol{\mu}'_\beta (\boldsymbol{\Sigma}_\beta)^{-1} \tag{A.5}$$

$$\boldsymbol{A} \equiv \sum \left(\mathbf{x}_t'(\sigma^2)^{-1}\mathbf{x}_t\right) + (\boldsymbol{\Sigma}_\beta)^{-1} \tag{A.6}$$

For the independent and identically-distributed error variance σ^2:

$$[\sigma^2|\cdot] \propto \left(\prod_{t=1}^{T}[\mathrm{y}_t|\mathbf{x}_t'\boldsymbol{\beta}, \sigma^2]\right)[\sigma^2] \tag{A.7}$$

Combine like terms and simplify to get an Inverse Gamma distribution with the following hyperparameters:

$$[\sigma^2|\cdot] \propto (\sigma^2)^{-(\tilde{q_y}+1)}\exp\left(\frac{-1}{\sigma^2}\frac{1}{\tilde{r_y}}\right) \tag{A.8}$$

$$\tilde{q_y} \equiv \frac{T}{2} + q_y \tag{A.9}$$

$$\tilde{r_y} \equiv \left(\frac{\sum_{t=1}^{T}(\mathrm{y}_t - \mathbf{x}_t'\boldsymbol{\beta})'(\mathrm{y}_t - \mathbf{x}_t'\boldsymbol{\beta})}{2} + \frac{1}{r_y}\right)^{-1} \tag{A.10}$$

A.2.2 Model 2

The full conditional statements are presented when the data are observed. The time series model is similar to the spatial model but different in a few key ways. The time series model is applied to each subregion separately (e.g., running the model once for each subregion) in a divide-and-conquer approach.

For the autocorrelated error component z_t at time t:

$$[z_t|\cdot] \propto \left(\prod_{t=1}^{T}[\mathrm{y}_t|\mathbf{x}_t'\boldsymbol{\beta}, z_t, \sigma_y^2][\mathrm{z}_t|\alpha, z_{t-1}, \sigma_z^2][\mathrm{z}_{t+1}|\alpha, z_t, \sigma_z^2]\right) \tag{A.11}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \sum \left((\mathrm{y}_t - \mathbf{x}_t'\boldsymbol{\beta})'(\sigma_y^2)^{-1} + (\alpha \mathrm{z}_{t-1})'(\sigma_z^2)^{-1} + \mathrm{z}_{t+1}'(\sigma_z^2)^{-1}\alpha \right) \tag{A.12}$$

$$\boldsymbol{A} \equiv \sum \left((\sigma_y^2)^{-1} + (\sigma_z^2)^{-1} + \alpha'(\sigma_z^2)^{-1}\alpha \right) \tag{A.13}$$

For the regression coefficients $\boldsymbol{\beta}$:

$$[\boldsymbol{\beta}|\cdot] \propto \left(\prod_{t=1}^{T} [\mathrm{y}_t | \mathbf{x}_t'\boldsymbol{\beta}, \mathrm{z}_t, \sigma_y^2] \right) [\boldsymbol{\beta}] \tag{A.14}$$

Complete the square and combine like terms to get a multivariate normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \sum \left((\mathrm{y}_t - \mathrm{z}_t)'(\sigma_y^2)^{-1}\mathbf{x}_t' \right) + \boldsymbol{\mu}_\beta'(\boldsymbol{\Sigma}_\beta)^{-1} \tag{A.15}$$

$$\boldsymbol{A} \equiv \sum \left(\mathbf{x}_t'(\sigma_y^2)^{-1}\mathbf{x}_t \right) + (\boldsymbol{\Sigma}_\beta)^{-1} \tag{A.16}$$

For the autocorrelation coefficient of order one α_1:

$$[\alpha|\cdot] \propto \left(\prod_{t=1}^{T} [\mathrm{z}_t | \alpha \mathrm{z}_{t-1}, \sigma_z^2] \right) [\alpha] \tag{A.17}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \sum \left(\mathrm{z}_t'(\sigma_z^2)^{-1}\mathrm{z}_{t-1} \right) + \mu_\alpha'(\sigma_\alpha^2)^{-1} \tag{A.18}$$

$$\boldsymbol{A} \equiv \sum \left(\mathrm{z}_{t-1}'(\sigma_z^2)^{-1}\mathrm{z}_{t-1} \right) + (\sigma_\alpha^2)^{-1} \tag{A.19}$$

For the independent and identically distributed variance σ_y^2:

$$[\sigma_y^2|\cdot] \propto \left(\prod_{t=1}^{T}[\mathrm{y}_t|\mathbf{x}_t'\boldsymbol{\beta}, \mathrm{z}_t, \sigma_y^2]\right)[\sigma_y^2] \tag{A.20}$$

Combine like terms and simplify to get an Inverse Gamma distribution with the following hyperparameters:

$$\tilde{q_y} \equiv \frac{T - NAs}{2} + q_y \tag{A.21}$$

$$\tilde{r_y} \equiv \left(\frac{\sum_{t=1}^{T}(\mathrm{y}_t - \mathbf{x}_t'\boldsymbol{\beta} - \mathrm{z}_t)'(\mathrm{y}_t - \mathbf{x}_t'\boldsymbol{\beta} - \mathrm{z}_t)}{2} + \frac{1}{r_y}\right)^{-1} \tag{A.22}$$

For the AR(1) variance σ_z^2:

$$[\sigma_z^2|\cdot] \propto \left(\prod_{t=1}^{T}[\mathrm{z}_t|\alpha, \mathrm{z}_{t-1}, \mathrm{z}_t, \sigma_z^2]\right)[\sigma_z^2] \tag{A.23}$$

Combine like terms and simplify to get an Inverse Gamma distribution with the following hyperparameters:

$$\tilde{q_z} \equiv \frac{T - 1}{2} + q_z \tag{A.24}$$

$$\tilde{r_z} \equiv \left(\frac{\sum_{t=1}^{T}(\mathrm{z}_t - \alpha\mathrm{z}_{t-1})'(\mathrm{z}_t - \alpha\mathrm{z}_{t-1})}{2} + \frac{1}{r_z}\right)^{-1} \tag{A.25}$$

For the autocorrelated error component z_t at time T:

$$[\mathrm{z}_T|\cdot] \propto [\mathrm{y}_T|\mathbf{x}_T, \boldsymbol{\beta}, \mathrm{z}_T, \sigma_y^2][\mathrm{z}_T|\alpha\mathrm{z}_{T-1}, \sigma_z^2] \tag{A.26}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$\boldsymbol{b'} \equiv (\mathrm{y}_T - x_T'\boldsymbol{\beta})'(\sigma_y^2)^{-1} + \alpha\mathrm{z}_{T-1}(\sigma_z^2)^{-1} \tag{A.27}$$

$$\boldsymbol{A} \equiv (\sigma_y^2)^{-1} + (\sigma_z^2)^{-1} \tag{A.28}$$

For the autocorrelated error component z_t at the initial state:

$$[\mathrm{z}_0|\cdot] \propto [\mathrm{z}_1|\alpha, \mathrm{z}_0, \sigma_z^2][\mathrm{z}_0] \tag{A.29}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \mathrm{z}_1'(\sigma_z^2)^{-1}\alpha + \mu_{z_0}'(\sigma_{z_0}^2)^{-1} \tag{A.30}$$

$$\boldsymbol{A} \equiv \alpha'(\sigma_z^2)^{-1}\alpha + (\sigma_{z_0}^2)^{-1} \tag{A.31}$$

A.2.3 Model 3

The spatial regression model is a Gaussian process model in a hierarchical Bayesian framework. The spatial model is different from the time series model because the model only requires one simulation to examine the differences at all subregions as opposed to a divide-and-conquer approach. The setup assumes independence at each subregion and the delineation for each subregion is governed by the rainfall runoff model.

For the spatially-correlated error component $\boldsymbol{\eta}_d$ sampled per subregion:

$$[\boldsymbol{\eta}_d|\cdot] \propto \left(\prod_{n=1}^{N_d} [\mathbf{y}_d|\mathbf{X}_d\boldsymbol{\beta}, \boldsymbol{\eta}_d, \sigma_d^2][\boldsymbol{\eta}_d|\boldsymbol{\Sigma}(\sigma_{\eta_d}^2, \rho)] \right) \tag{A.32}$$

Complete the square and combine like terms to get a multivariate normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \sum \left((\mathbf{y}_d - \mathbf{X}_d'\boldsymbol{\beta})'(\sigma_d^2)^{-1} \right) + \boldsymbol{\mu}_\eta'\boldsymbol{\Sigma}(\sigma_{\eta_d}^2, \rho)^{-1} \tag{A.33}$$

$$\boldsymbol{A} \equiv \sum \left((\sigma_d^2 \mathbf{I})^{-1} + \boldsymbol{\Sigma}(\sigma_{\eta_d}^2, \rho)^{-1} \right) \tag{A.34}$$

For the regression coefficient vector $\boldsymbol{\beta}$ sampled *across* all subregions:

$$[\boldsymbol{\beta}|\cdot] \propto \left(\prod_{d=1}^{D} [\mathbf{y}_d|\mathbf{X}_d\boldsymbol{\beta}, \boldsymbol{\eta}_d, \sigma_d^2] \right) [\boldsymbol{\beta}] \tag{A.35}$$

Complete the square and combine like terms to get a multivariate normal distribution with the following hyperparameters:

$$\boldsymbol{b}' \equiv \sum \left((\mathbf{y}_d - \boldsymbol{\eta}_d)'(\sigma_d^2)^{-1}\mathbf{X}_d \right) + \boldsymbol{\mu}_\beta' \boldsymbol{\Sigma}_\beta^{-1} \tag{A.36}$$

$$\boldsymbol{A} \equiv \sum \left(\mathbf{X}_d'(\sigma_d^2)^{-1}\mathbf{X}_d \right) + \boldsymbol{\Sigma}_\beta^{-1} \tag{A.37}$$

For the independent and identically-distributed variance σ_d^2 sampled for per subregion:

$$[\sigma_d^2|\cdot] \propto \left(\prod_{n=1}^{N_d} [\mathbf{y}_d|\mathbf{X}_d'\boldsymbol{\beta}, \boldsymbol{\eta}_d, \sigma_d^2] \right) [\sigma_d^2] \tag{A.38}$$

Combine like terms to simplify:

$$[\sigma_d^2|\cdot] \propto (\sigma_d^2)^{-(\tilde{q_d}+1)} \exp\left(\frac{-1}{\sigma_d^2} \frac{1}{\tilde{r_d}} \right) \tag{A.39}$$

To get an Inverse Gamma distribution with the following hyperparameters:

$$\tilde{q_d} \equiv \frac{n_d}{2} + q_d \tag{A.40}$$

$$\tilde{r_d} \equiv \left(\frac{\sum_{n=1}^{N_d} (\mathbf{y}_d - \mathbf{X}_d'\boldsymbol{\beta} - \boldsymbol{\eta}_d)'(\mathbf{y}_d - \mathbf{X}_d'\boldsymbol{\beta} - \boldsymbol{\eta}_d)}{2} + \frac{1}{r_d} \right)^{-1} \tag{A.41}$$

For the spatially-correlated variance component $\sigma_{\eta_d}^2$ sampled per subregion:

$$[\sigma^2_{\eta_d}|\cdot] \propto \left(\prod_{n=1}^{N_d}[\boldsymbol{\eta}_d|\boldsymbol{\Sigma}(\sigma^2_{\eta_d},\rho)^{-1}]\right)[\sigma^2_{\eta_d}] \tag{A.42}$$

Combine like terms to simplify:

$$[\sigma^2_{\eta_d}|\cdot] \propto (\sigma^2_{\eta_d})^{-(\tilde{q_\eta}+1)}\exp\left(\frac{-1}{\sigma^2_{\eta_d}}\frac{1}{\tilde{r_d}}\right) \tag{A.43}$$

To get an Inverse Gamma distribution with the following hyperparameters:

$$\tilde{q_\eta} \equiv \frac{n_\eta}{2} + q_\eta \tag{A.44}$$

$$\tilde{r_\eta} \equiv \left(\frac{\sum_{n=1}^{N_d}(\boldsymbol{\eta}_d - \boldsymbol{\mu}_\eta)'(\mathbf{R}(\rho))^{-1}(\boldsymbol{\eta}_d - \boldsymbol{\mu}_\eta)}{2} + \frac{1}{r_\eta}\right)^{-1} \tag{A.45}$$

For the spatial correlation coefficient ρ (similar to α in the time series analysis) sampled *across* all subregions:

$$[\rho|\cdot] \propto \left(\prod_{d=1}^{D}[\boldsymbol{\eta}_d|\boldsymbol{\Sigma}(\sigma^2_{\eta_d},\rho)^{-1}]\right)[\rho] \tag{A.46}$$

which is not conjugate so an M-H (Metropolis-Hastings) update is required by sampling ρ^* from the prior as a proposal and using the M-H ratio:

$$\text{mh} = \frac{[\boldsymbol{\eta}_d|\boldsymbol{\Sigma}(\sigma^{2(k)}_{\eta_d},\rho^*)^{-1}]}{[\boldsymbol{\eta}_d|\boldsymbol{\Sigma}(\sigma^{2(k)}_{\eta_d},\rho^{k-1})^{-1}]} \tag{A.47}$$

A.3 Tables

Table A.1: A summary of the groundwater observation stations used in this study. All groundwater observations are measured by USGS. The station IDs are defined by CO-DWR.

Subregion	USGS ID	Lat/Long	Record	Dist. to River
1	380443102555200	38.08°, -102.93°	09/1958 - 11/2018	800 m
2	380450102464200	38.08°, -102.78°	10/1963 - 11/2018	1200 m
3	380450102384000	38.08°, -102.65°	08/1965 - 11/2018	2500 m
4	380648102342900	38.11°, -102.58°	10/1953 - 11/2018	70 m
5	380651102285900	38.11°, -102.48°	04/1964 - 11/2018	250 m
6	380654102245602	38.11°, -102.42°	01/1967 - 11/2018	975 m
7	380553102183800	38.10°, -102.31°	04/1964 - 11/2018	215 m
8	380418102112700	38.07°, -102.19°	06/1965 - 11/2018	3000 m

Table A.2: A summary of confluences with the main stem of the Arkansas River. Limestone Creek and Graveyard Creek are currently monitored by Colorado Division of Water Resource (DWR) with records that begin 02/2017 and 10/2016, respectively; Big Sandy Creek and Wild Horse Creek are currently monitored by USGS with records that begin in 01/1999 and 03/1999, respectively. Based on the USGS topographical map, Buffalo Creek and Two Butte Creek are both considered to be an intermittent wash. On CDSS's Map Viewer, West Prowers Arroyo is also named Levere Ditch and East Prowers Arroyo is named Lubers Drainage Ditch. Keesee and Fort Bent canals are augmented which impact subregion two. Keesee, Fort Bent and Amity canals are augmented which impact subregion three. Lamar and X-Y Graham canals are augmented which impact subregion five and six. Sisson-Stubbs canal is augmented which impacts subregion eight. Subregion three is impacted by pastoral practices. Subregion four is impacted by the city of Lamar. Subregion seven and eight are impacted by the city of Granada and Holly, respectively. Subregion six is impacted by storage ponds adjacent to a percentage of the riparian area.

Stream Name	Lat/Long	Dist. (km)	Stream Type
Caddoa Creek	38.08°, -102.90°	94.8	Intermittent
Limestone Creek	38.08°, -102.86°	88.7	Intermittent
Mud Creek	38.08°, -102.84°	86.5	Perennial
Graveyard Creek	38.09°, -102.79°	80.9	Intermittent
Clay Creek	38.10°, -102.53°	51.5	Intermittent
Big Sandy Creek	38.11°, -102.48°	45.3	Perennial
Cottonwood Creek	38.11°, -102.39°	35.0	Intermittent
Boggs Creek	38.11°, -102.37°	34.0	Intermittent
Wolf Creek	38.09°, -102.31°	27.2	Intermittent
Buffalo Creek	38.09°, -102.29°	25.1	Intermittent
Wild Horse Creek	38.05°, -102.13°	8.7	Perennial
Two Butte Creek	38.04°, -102.13°	7.9	Intermittent
Cheyenne Creek	38.03°, -102.05°	0.35	Intermittent
West Prowers Arroyo	38.08°, -102.92°	96.6	Ephemeral
East Prowers Arroyo	38.09°, -102.88°	91.5	Ephemeral
McClave Drainage Ditch	38.09°, -102.82°	83.9	NA
Riverview Ditch	38.10°, -102.73°	75.2	NA
Wiley Drainage Ditch	38.10°, -102.70°	71.9	NA
Markham Arroyo	38.11°, -102.58°	58.1	Ephemeral
Fort Lyon	38.01°, -103.59°	NA	Canal
Fort Bent	38.08°, -102.86°	89.7	Canal
Keesee	38.08°, -102.86°	89.7	Canal
Amity	38.09°, -102.76°	78.2	Canal
Lamar	38.11°, -102.64°	64.7	Canal
Hyde	38.11°, -102.61°	61.7	Canal
Manvel	38.10°, -102.49°	47.3	Canal
X-Y Graham	38.10°, -102.43°	39.2	Canal
Buffalo	38.10°, -102.33°	29.1	Canal
Sisson-Stubbs	38.05°, -102.22°	16.6	Canal

Table A.3: A summary of the canals used in this study. The asterisk next to Fort Lyon, Lamar, Amity, and Fort Bent canals indicate that each canal has seasonally-augmented flows but still divert water for surface irrigation. The River Distance statistic is recorded from CDSS and converted to kilometers.

Name	Dist. (km)	Record	Augment?	Subregion
Fort Lyon	NA	11/1910-pres	Yes*	1 and 2
Fort Bent	89.7	11/1938-pres	Yes*	2 and 3
Keesee	89.7	11/1938-03/2003	Yes	2 and 3
Amity	78.2	11/1938-pres	Yes*	3, 7 and 8
Lamar	64.7	11/1938-pres	Yes*	4 and 5
Hyde	61.7	11/1938-pres	No	4
Manvel	47.3	11/1938-9/1993	Yes	5
X-Y Graham	39.2	8/1939-12/1995	Yes	6 and 7
Buffalo	29.1	11/1938-pres	No	7 and 8
Sisson-Stubbs	16.6	7/1948-10/1982	Yes	8

Appendix B

Chapter 4: Supplementary Material

B.1 Model Posteriors

B.1.1 Control Model Posterior

$$[\beta_0, \beta_1, \sigma_y^2|\mathrm{y}_i] \propto \left(\prod_{n=i}^{N}[\mathrm{y}_i|\beta_0, \beta_1, \sigma_y^2]\right)[\sigma_y^2][\beta_0][\beta_1] \tag{B.1}$$

B.1.2 Asymptote Model Posterior

$$[a, v_{max}, c, \sigma_y^2|\mathrm{y}_i] \propto \left(\prod_{n=i}^{N}[\mathrm{y}_i|a, c, v_{max}, \sigma_y^2]\right)[\sigma_y^2][a][c][v_{max}] \tag{B.2}$$

B.1.3 Change Point Model Posterior

$$[\beta_0, \beta_1, \beta_2, cp, \sigma_y^2|\mathrm{y}_i] \propto \left(\prod_{n=i}^{N}[\mathrm{y}_i|\beta_0, \beta_1, \beta_2, cp, \sigma_y^2]\right)[\sigma_y^2][\beta_0][\beta_1][\beta_2][cp] \tag{B.3}$$

B.2 Full Conditional Statements

B.2.1 Control Model

For the intercept regression coefficient β_0:

$$[\beta_0|\cdot] \propto \left(\prod_{i=1}^{N}[\mathrm{y}_t|\beta_0, \sigma^2]\right)[\beta_0] \tag{B.4}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$b \equiv \sum\left((\mathrm{y}_i - \beta_1\mathrm{x}_{i,2})\mathrm{x}_{i,1}\right)(\sigma_y^2)^{-1} + \mu_{\beta_0}(\sigma_{\beta_0}^2)^{-1} \tag{B.5}$$

$$a \equiv \sum \left(\mathrm{x}_{i,1}\right)^2 \left(\sigma_y^2\right)^{-1} + \left(\sigma_{\beta_0}^2\right)^{-1} \tag{B.6}$$

For the slope regression coefficient β_1:

$$[\beta_1|\cdot] \propto \left(\prod_{i=1}^{N} [\mathrm{y}_t|\beta_1, \sigma_y^2]\right) [\beta_1] \tag{B.7}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$b \equiv \sum \left((\mathrm{y}_i - \beta_0 \mathrm{x}_{i,1})\mathrm{x}_{i,2}\right) \left(\sigma_y^2\right)^{-1} + \mu_{\beta_1} (\sigma_{\beta_1}^2)^{-1} \tag{B.8}$$

$$a \equiv \sum \left(\mathrm{x}_{i,2}\right)^2 \left(\sigma_y^2\right)^{-1} + \left(\sigma_{\beta_1}^2\right)^{-1} \tag{B.9}$$

For the independent and identically-distributed error variance σ_y^2:

$$[\sigma_y^2|\cdot] \propto \left(\prod_{i=1}^{N} [\mathrm{y}_i|g(x)_i, \sigma_y^2]\right) [\sigma_y^2] \tag{B.10}$$

Combine like terms and simplify to get an Inverse Gamma distribution with the following hyperparameters:

$$[\sigma_y^2|\cdot] \propto \left(\sigma_y^2\right)^{-(\tilde{q_y}+1)} \exp\left(\frac{-1}{\sigma_y^2}\frac{1}{\tilde{r_y}}\right) \tag{B.11}$$

$$\tilde{q_y} \equiv \frac{N}{2} + q_y \tag{B.12}$$

$$\tilde{r_y} \equiv \left(\frac{\sum_{i=1}^{N} (\mathrm{y}_i - g(x)_i)^2}{2} + \frac{1}{r_y}\right)^{-1} \tag{B.13}$$

B.2.2 Asymptote Model

For the v_{max} parameter:

$$[\mathrm{v}_{max}|\cdot] \propto \left(\prod_{i=1}^{N}[\mathrm{y}_i|\mathrm{v}_{max}, \sigma_y^2)]\right)[\mathrm{v}_{max}] \tag{B.14}$$

which is not conjugate so an M-H (Metropolis-Hastings) update is required by sampling v^*_{max} from a symmetric proposal, updating $\mathbf{f(x)}^{(*)} = \frac{\mathrm{v}^{(*)}_{max}\times(\mathbf{x}+c)}{\frac{\mathrm{v}^{(*)}_{max}}{a}+(\mathbf{x}+c)}$, if v^*_{max} is negative then skip computing the mh ratio and let $\mathrm{v}^{(k)}_{max} = \mathrm{v}^{(k-1)}_{max}$, otherwise compute the the M-H ratio:

$$\mathrm{mh} = \frac{\left[\mathbf{y} \mid \mathbf{f(x)}^{(*)}, \sigma_y^{2,(k)}\right]\left[\mathrm{v}^{(*)}_{max}\right]}{\left[\mathbf{y} \mid \mathbf{f(x)}^{(k-1)}, \sigma_y^{2,(k)}\right]\left[\mathrm{v}^{(k-1)}_{max}\right]} \tag{B.15}$$

For the a parameter:

$$[a|\cdot] \propto \left(\prod_{i=1}^{N}[\mathrm{y}_i|a, \sigma_y^2)]\right)[a] \tag{B.16}$$

which is not conjugate so an M-H (Metropolis-Hastings) update is required by sampling a^* from a symmetric proposal, updating $\mathbf{f(x)}^{(*)} = \frac{\mathrm{v}_{max}\times(\mathbf{x}+c)}{\frac{\mathrm{v}_{max}}{a^{(*)}}+(\mathbf{x}+c)}$, if a^* is negative then skip computing the mh ratio and let $a^{(k)} = a^{(k-1)}$, otherwise compute the the M-H ratio:

$$\mathrm{mh} = \frac{\left[\mathbf{y} \mid \mathbf{f(x)}^{(*)}, \sigma_y^{2,(k)}\right]\left[a^{(*)}\right]}{\left[\mathbf{y} \mid \mathbf{f(x)}^{(k-1)}, \sigma_y^{2,(k)}\right]\left[a^{(k-1)}\right]} \tag{B.17}$$

For the c parameter:

$$[c|\cdot] \propto \left(\prod_{i=1}^{N}[\mathrm{y}_i|c, \sigma_y^2)]\right)[c] \tag{B.18}$$

which is not conjugate so an M-H (Metropolis-Hastings) update is required by sampling c^* from a symmetric proposal, updating $\mathbf{f(x)}^{(*)} = \frac{\mathrm{v}_{max}\times(\mathbf{x}+c^*)}{\frac{\mathrm{v}_{max}}{a}+(\mathbf{x}+c^*)}$ and compute the the M-H ratio:

$$\text{mh} = \frac{\left[\mathbf{y} \mid \mathbf{f(x)}^{(*)}, \sigma_y^{2,(k)}\right]\left[c^{(*)}\right]}{\left[\mathbf{y} \mid \mathbf{f(x)}^{(k-1)}, \sigma_y^{2,(k)}\right]\left[c^{(k-1)}\right]} \tag{B.19}$$

For the independent and identically-distributed error variance σ_y^2:

$$[\sigma_y^2|\cdot] \propto \left(\prod_{i=1}^{N}[\mathrm{y}_i|g(x)_i, \sigma_y^2]\right)[\sigma_y^2] \tag{B.20}$$

Combine like terms and simplify to get an Inverse Gamma distribution with the following hyperparameters:

$$[\sigma_y^2|\cdot] \propto (\sigma_y^2)^{-(\tilde{q_y}+1)}\exp\left(\frac{-1}{\sigma_y^2}\frac{1}{\tilde{r_y}}\right) \tag{B.21}$$

$$\tilde{q_y} \equiv \frac{N}{2} + q_y \tag{B.22}$$

$$\tilde{r_y} \equiv \left(\frac{\sum_{i=1}^{N}(\mathrm{y}_i - g(x)_i)^2}{2} + \frac{1}{r_y}\right)^{-1} \tag{B.23}$$

B.2.3 Change Point Model

If the covariate vector is $\mathrm{x}_{i,1}$ and $\mathrm{x}_{i,2}$ is a dummy vector where $\mathrm{x}_{i,2} = 0$ when $\mathrm{x}_{i,2} \leq cp$ and 1 if $\mathrm{x}_{i,2} > cp$, then for the change point parameter:

$$[cp|\cdot] \propto \left(\prod_{i=1}^{N}[\mathrm{y}_t|cp, \sigma_y^2]\right)[cp] \tag{B.24}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$b \equiv \sum\Big((\mathrm{y}_i - \beta_0 - \beta_1\mathrm{x}_{i,1} - \beta_2\mathrm{x}_{i,1}\mathrm{x}_{i,2})((-1)\beta_2\mathrm{x}_{i,2})\Big)(\sigma_y^2)^{-1} + \mu_{cp}(\sigma_{cp}^2)^{-1} \tag{B.25}$$

$$a \equiv \sum \Big(((-1)\beta_2 \mathrm{x}_{i,2})^2 \Big) (\sigma_y^2)^{-1} + (\sigma_{cp}^2)^{-1} \tag{B.26}$$

For the intercept regression coefficient β_0:

$$[\beta_0|\cdot] \propto \left(\prod_{i=1}^{N} [\mathrm{y}_i|\beta_0, \sigma^2] \right) [\beta_0] \tag{B.27}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$b \equiv \sum \Big((\mathrm{y}_i - \beta_1 \mathrm{x}_{i,1} - \beta_2 (\mathrm{x}_{i,1} - \mathrm{cp}) \mathrm{x}_{i,2}) \Big) (\sigma_y^2)^{-1} + \mu_{\beta_0} (\sigma_{\beta_0}^2)^{-1} \tag{B.28}$$

$$a \equiv \sum \Big(1 \Big) (\sigma_y^2)^{-1} + (\sigma_{\beta_0}^2)^{-1} \tag{B.29}$$

For the first slope regression coefficient β_1:

$$[\beta_1|\cdot] \propto \left(\prod_{i=1}^{N} [\mathrm{y}_i|\beta_1, \sigma^2] \right) [\beta_1] \tag{B.30}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$b \equiv \sum \Big((\mathrm{y}_i - \beta_0 - \beta_2 (\mathrm{x}_{i,1} - \mathrm{cp}) \mathrm{x}_{i,2}) \mathrm{x}_{i,1} \Big) (\sigma_{y,1}^2)^{-1} + \mu_{\beta_1} (\sigma_{\beta_1}^2)^{-1} \tag{B.31}$$

$$a \equiv \sum \Big(\mathbf{x}_{i,1}^2 \Big) (\sigma_y^2)^{-1} + (\sigma_{\beta_1}^2)^{-1} \tag{B.32}$$

For the second slope regression coefficient β_2:

$$[\beta_2|\cdot] \propto \left(\prod_{i=1}^{N} [\mathrm{y}_i|\beta_2, \sigma^2] \right) [\beta_2] \tag{B.33}$$

Complete the square and combine like terms to get a normal distribution with the following hyperparameters:

$$b \equiv \sum \Big((\mathrm{y}_i - \beta_0 - \beta_1 \mathrm{x}_{i,1})(\mathrm{x}_{i,1}\mathrm{x}_{i,2} - \mathrm{cpx}_{i,2}) \Big) (\sigma_y^2)^{-1} + \mu_{\beta_2} (\sigma_{\beta_2}^2)^{-1} \tag{B.34}$$

$$a \equiv \sum \Big((\mathrm{x}_{i,1}\mathrm{x}_{i,2} - \mathrm{cpx}_{i,2})^2 \Big) (\sigma_y^2)^{-1} + (\sigma_{\beta_2}^2)^{-1} \tag{B.35}$$

For the independent and identically-distributed error variance σ_y^2:

$$[\sigma_y^2|\cdot] \propto \left(\prod_{i=1}^{N} [\mathrm{y}_i | g(x)_i, \sigma_y^2] \right) [\sigma_y^2] \tag{B.36}$$

Combine like terms and simplify to get an Inverse Gamma distribution with the following hyperparameters:

$$\widetilde{q}_y = \frac{N}{2} + q_y \tag{B.37}$$

$$\widetilde{r}_y = \left(\frac{\sum_{i=1}^{N} (y_i - (\beta_0 + \beta_1 \mathrm{x}_{1,i} + \beta_2 (\mathrm{x}_{1,i} - \mathrm{cp}) \mathrm{x}_{1,2}))^2}{2} + \frac{1}{r_y} \right)^{-1} \tag{B.38}$$